THE UNIVERSITY OF ALBERTA

A STATISTICAL ANALYSIS OF THE PROFILE TECHNIQUE FOR THE EVALUATION OF COMPETITIVE BASKETBALL PERFORMANCE

by

H. KENNER KAY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF ARTS

FACULTY OF PHYSICAL EDUCATION

EDMONTON, ALBERTA

AUGUST, 1966

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled " A Statistical Analysis Of The Profile Technique For The Evaluation Of Competitive Basketball Performance", submitted by H. Kenner Kay in partial fulfilment of the requirements for the degree of Master of Arts.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

STATEMENT OF THE PROBLEM

Introduction. Evaluation by means of appropriate tests is a valuable asset to the teacher of physical education in preparing and conducting an efficient physical education program (6). Progress in physical education toward the attainment of objectives is dependent upon the quality of its research and the validity of its measurement procedures (6).

The majority of basketball tests that have been proposed are based on opinions of the authors and are generally unsubstantiated by scientific evidence (6). Virtually all of the existing basketball tests are standardized skills tests for non-competitive situations. Recent research has pointed out what many basketball coaches have considered true for many years. Motor performance is extremely specific (8,21). It would seem from the research of Triplet, Berridge and Gordon (8) that basketball skill performance in a series of non-competitive isolated drills is a completely different thing from skill performance in a competitive, integrated, game situation. As Branch McCracken (31) emphasizes that many players shoot well in warm up but perform poorly in game situations. Both positive and negative performance effects would result in the competitive situation depending on individual differences (8).

Statement of the Problem. There is no basketball test presently available which measures basketball performance in a competitive game situation and which has been substantiated by scientific evidence to meet the requisites listed by Clarke (6) as criteria for a test. It was the purpose of the present study to develop an objective instrument which would be valid, reliable and economical in terms of cost of the in-

struments and time for the evaluation of basketball performance in a game situation.

Sub-Problems. Secondary purposes of the study were: (1) to determine the length of performance necessary to obtain valid and reliable evaluation; (2) to determine if the instrument was suitable for evaluation of: intercollegiate varsity basketball, senior high school varsity basketball, and second year high school class basketball; (3) to determine if the amount of training necessary for an observer to obtain accurate results; (5) to determine whether reliability and item validity could be improved by changing the point ratio of the Profile technique.

Importance of the Study. Clarke (6) emphasizes "Progress in physical education toward the attainment of objectives is dependent upon the quality of its research and the validity of its measurement procedures." He believes that measurement is indispensable to an adequate physical education program (6). Measurement aids the physical educator to plan the direction and emphasis of his program. It provides him with a basis upon which to decide if the aims and objectives of the physical education or athletic program have been achieved and to plan for fuller relization of these goals in the future.

Appropriate measurement enables the instructor to determine pupil progress which provides the basis for individual instruction. Realistic tests provide teachers and students with a diagnostic device for the evaluation of strengths and weaknesses which under normal circumstances will give direction and motivation for future efforts.

Kenyon (25) states, "... the use of measurement in no way disregards the art in teaching physical education. Rather it provides some scientific material with which the teacher may improve his art."

Existing tests are of the non-competitive standardized skill type. Some of these tests are quite reliable and perhaps valid for measuring those specific skills sampled in the non-competitive situation, grading in a beginning basketball unit where the development of skills and not the development of basketball playing ability is the objective and to match students with others of near equal ability for the practising of skills and competition. But the work done on specificity by Henry (21) and others as summarized by Cratty (8) would imply that standardized skill tests are insufficient to measure basketball performance in a game situation.

Many of the existing tests fail to discriminate in a narrow range. Clarke (6) reports that the reliability of a number of the tests like the Johnson Test and the Knox Basketball Test as researched by Boyd, McCachren and Waglow (6) were found by the biserial correlation technique. This technique only discriminates into dichotomous categories - passed or failed, have or have not, basketball player or not a basketball player (16). It would seem that basketball is not a dichotomous ability but rather a continuum.

There is a great need for a test which in addition to meeting the basic criteria of a "good test" (6) will measure performance in the skills of basketball as they are used in the competitive situation and will be diagnostic in revealing the source of individual errors and the weakness of the basketball program in developing certain skills.

Delimitations of the Study. The study of the intercollegiate basketball was confined to the varsity teams of the University of Alberta at Edmonton (six games), University of Manitoba (two games), University

of Saskatchewan (two games), University of Alberta at Calgary (four games), University of Windsor (two games), Acadia University (two games), and Carleton University (two games). The study of high school varsity basketball was confined to the following teams from the Edmonton, Alberta high school senior league: Strathcona (three games), McNally (one game), and Bonnydoon (one game). The high school class study was confined to four games involving two classes of grade ten boys from Strathcona Composite High School in Edmonton Alberta after their second season of formal basketball instruction. The entire study was limited to the evaluation of male subjects.

Limitations of the Study. The principal limitation of the study is that the Profile Basketball Test yields scores that are relative to the level of competition. Nocomparison can be made in scores between groups at various levels of performance. Intercollegiate scores can not be compared with high school varisty scores. And varsity scores can not be compared with high school class scores. Because of this, no norms can be developed for this test. Only those factors which can be objectively scored in a quantitative form are taken into consideration in the evaluation.

In this study reliability of the technique is of essence rather than reliability of player performance. It is well known that basketball performance varies from moment to moment as the physiological and psychological state of the performer changes due to warm up, fatigue, competitive and speed factors. It is thought that split half reliability will minimize these factors.

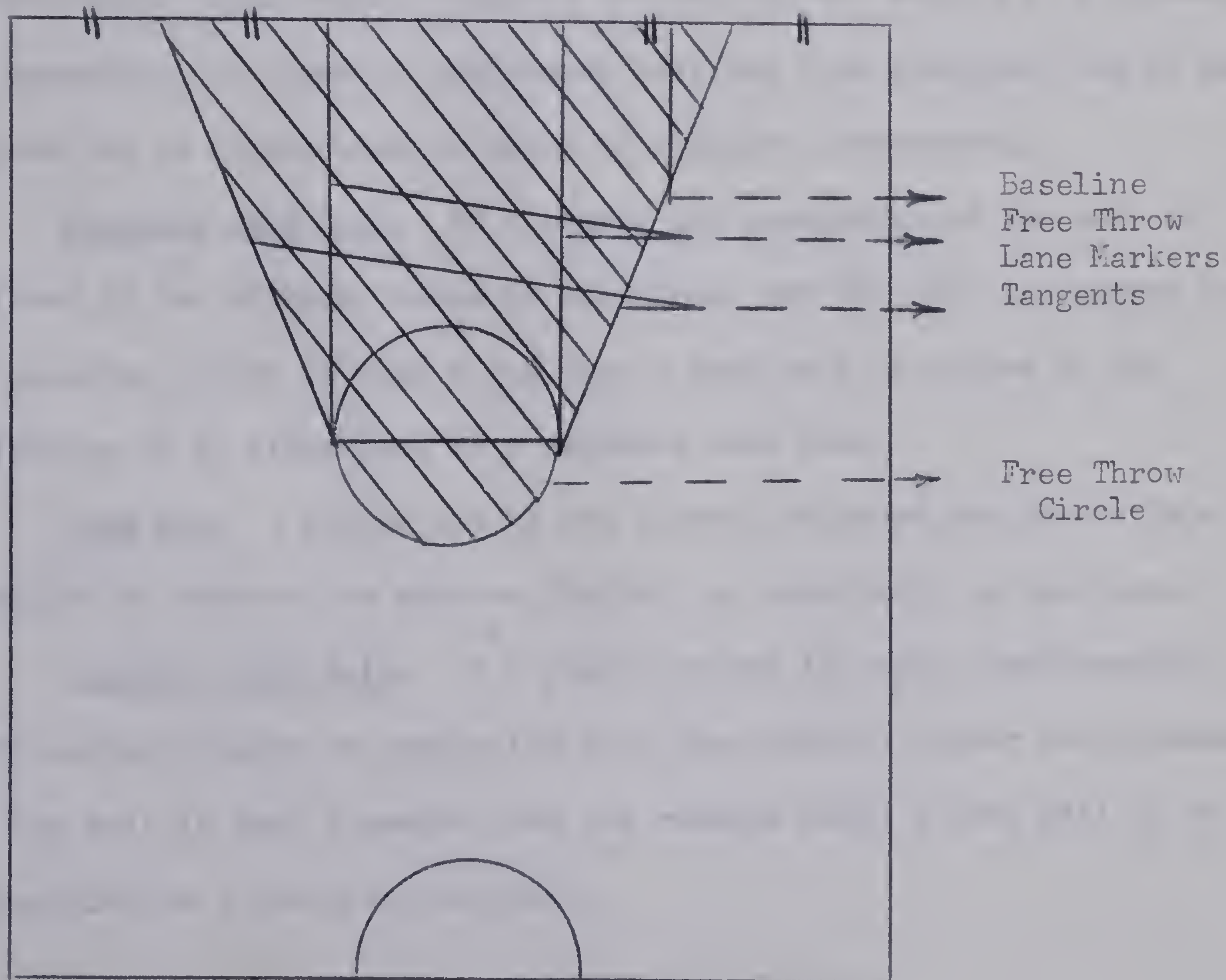
In varsity game situations time played and type of opposition were not directly controlled. The effect of these variables will have

to be determined by comparison of results with other parts of the study in which they were controlled and by partialling out the time played by each player.

DEFINITION OF TERMS USED

<u>Ability</u>. Ability is the possession of a relatively enduring competence or skill. In this study the profile technique will be operationally considered to be a reliable measure of basketball ability if a reliability coefficient significant at the .01 level of confidence for performance over a number of occassions is obtained.

<u>Assist</u>. An assist is any act which gives a teammate an unhindered shot within the area defined by the freethrow circle and the tangents to it from the points half way between the corner and the places where the free throw lane bisects the baseline.



ASSIST AREA

Bad Pass. Any pass to which the receiver is not able to get his full palm on the ball, or can not place his full palm on the ball with one step and a reach, or which is deflected in any way and which is not caught by the intended receiver is a bad pass,

Blocking Out. Blocking out is the act of legally positioning oneself so as to obstruct an opponent from reaching a rebound.

Free Ball. Any time the ball is not clearly in possession of one team or the other according to the official rules of basketball, it is classified as a free ball.

Fumble. Any time a player has lost possession of the ball after his team has had control of it, constitutes a fumble unless it falls under the definition of a bad pass, violation, foul, or missed shot.

Good Test. A test shall be considered a "good test" if it can be administered efficiently, measures the quality for which it is intended, is economical in terms of instrument cost and time involved, and if test scores can be interpreted in terms of relative performance.

Negative Held Ball. If a player has possession of the ball as defined by the official rules of basketball and the ball is grasped by an opposing player in such a way that a jump ball is called by the officials it is classified as a negative held ball.

Open Man. A player who is not closely defended and who is in a position to score or to advance the ball is considered an open man.

Positive Held Ball. If a player grasps the ball simultaneously with another player or grasps the ball when another player has possession of the ball in such a manner that the referee calls a jump ball it is classified as a positive held ball.

Performance. Performance is execution during a single occasion.

Rebound. If a player gains control of the ball after it has hit the rim or backboard and before it hits the floor or is touched by another player, it is classified as a rebound.

Recovery. A recovery shall be operationally defined as the act of obtaining possession of the ball in any way other than that defined as a rebound after it has been a free ball or in possession of the oppossing team as defined by the official rules of basketball.

Reliability. In this study reliability is defined as the degree of consistency with which the profile technique accurately measures basketball performance. This quality will be very difficult to determine because individual performance varies from moment to moment as the physiological and psychological state of the performers change due to differential warm up, fatigue, competitive and speed factors. It is considered that the split half method of determining reliability will minimize the effects of performance fluctuations and best show the reliability of the technique.

Stop Time. The method of controlling basketball playing time in which the official time keeper's clock is stopped when the ball is out of play according to the official rules of basketball is known as stop time.

Straight Time. Straight time is the method of controlling basketball playing time in which the official time keeper starts the time clock at the beginning of the playing period and does not stop the clock until the playing period is completed.

Violation. Any act which is defined as a violation according to the officail rules of basketball and which is called by the referee is scored as a violation.

CHAPTER II

REVIEW OF THE LITERATURE

Background Information. Basketball is a comparatively modern team game introduced by Dr. James A. Naismith in 1892 while he was working as physical director of the Y.M.C.A. in Springfield, Massachussetts (2). The game quickly spread throughout the Eastern United States, and then, as schools realized its educational possibilities, it spread to the rest of the United States and throughout the world (22). In Canada, basketball has not achieved the position it has as a spectator sport in the United States but for the participant it is a major activity. Almost every high school extramural, intramural, and class program has basketball as an integral part. There is hardly a Y.M.C.A. or Community Recreation program that does not include basketball. Dean (10:33) reports McCracken as follows:

> Fundamentals are the backbone of any successful team. Boys who know their fundamentals and have the physical makeup with a desire to win will make any system look good .... The successful coach is the one who can sell the boys on the importance of fundamentals and good physical condition and importance of working together as a unit.

Cappan as quoted in Dean (10:38), also places the emphasis on fundamentals in his statement that basketball is a team game played by boys with highly developed individual skills. And Holman (22) states, "... mastery of detail in fundamentals is the first principle of winning basketball."

Different coaches and experts on basketball have slightly different views on the relative importance of various fundamentals in basketball but few, if any, will disagree that execution of fundamentals is paramount in basketball performance.

Mauer as quoted by Dean (10:41) states that possession of the ball means control of the game.

> Most coaches, when rating a player, first think in terms of ball-handling ability. The only substitute for this fundamental might be great shooting ability. The test of an offense is its ability to control the ball while working into good close-in shooting position.

Individual characteristics to consider in selecting basketball players are: ball-handling, shooting ability and players who play well under pressure in a game situation (42). It has been estimated that about fifty percent of all basketball games are won or lost at the free throw line (10). McCracken (31) feels that passing, shooting and gaining possession of the ball through "held balls", "rebounds", and recovery of "free balls" is so essential to basketball performance that he devoted separate chapters to each of these fundamentals. He states that the important things in evaluating an opponent are: shooting ability, area in which shots are taken, where they are most accurate, how much and how well they dribble, defensive ability, cleverness in stops and turns, rebounding ability, and foul shooting accuracy (31). When we think of a basketball player, there are many things to take into consideration, but the final question remains, "Is he a scoring threat?" (31:74) "There is a common saying, 'You can't score if you don't have the ball'. The same thing applies to your opposition. They can't score as long as you have the ball." (31:104)

There are only three ways of gaining possession of the ball: held balls, rebounds and "recovery" of free balls (31). The rebound on both the offensive and defensive back boards is a very important factor in the outcome of any basketball game (31). To retain possession take only good shots, throw accurate passes, do not be

forced into a held ball situation, and know where your teammates are to get the ball to an "open man" (31). An important point on defense is to assume and maintain good defensive position which will enable efficient "blocking out" for rebounds and will help the player be alert for interceptions.

Existing Basketball Tests. One of the earliest batteries of tests designed to measure basketball techniques was developed by Brace (4) in 1924. This battery included six tests: the single overarm target throw, two arm overhead target throw, two handed push pass, rapid pass, position dribbling, and dribble and shoot test. No statistical analysis of these tests was made.

Edgren (11) was the first to report a basketball test based upon statistical analysis. Eight items were used in this test including as many of the game skills as possible: speed pass, accuracy pass, pivot and shoot, speed dribble, dribble and shoot, accuracy shooting, opposition shooting and ball handling. A validity coefficient of .77 was obtained between the test battery and a subjective rating of the performance of players.

Some of the speed items of the Edgren test are quite reliable and have been adopted by more recent investigators into their tests (6,28,43). Cratty (8) points out that most motor skills show considerable performance fluctuations from trial to trial. It seems very unlikely that the five trials given in the accuracy pass, pivot and shoot, dribble and shoot, opposition shooting and the ten trials of the accuracy shooting test would be sufficient to cancel out the effect of performance fluctuations. The pivot and shoot test does not have any

control on the speed of shooting or the manner in which the pivot and shot are integrated.

Lehsten (28) developed a five item test including: baskets per minute, vertical jump, 40-foot dash, wall bounce and a dodging run. The Lehsten test must be considered a test of potential basketball ability rather than an ability test itself. Three of the five tests: the vertical jump, the 40-foot dash and the dodging run, do not involve either a basketball or skills specific to the game.

Young and Moser (43) developed a five item test. They adopted the Edgren ball handling test (11) and a wall speed test which involved passing the ball against the wall from behind a restraining line 6 feet from the wall as often as possible in 30 seconds. The moving target test involved throwing a basketball ten times at a target swinging on a pendulum. The bounce and shoot test involved five lay-up shots from a dribble on each side of the basket. It would seem that ten trials would be insufficient to cancel out the performance fluctuations spoken of by Cratty (8). The free jump is an elaborate form of vertical jump which might be one measure of potential basketball ability but is not a measure of basketball ability itself.

The Wisconsin basketball battery (18) consists of three items. The wall speed test was adopted from the Young-Moser Tests (43). The zone toss test involves throwing a ball over a rope 7 feet 1 inch from the floor and across a zone 6 feet 4 inches wide. Scoring is based on the time taken by a subject to make ten such tosses running from zone to zone under the rope to catch one's own throw. In view of the work done by Henry (21) and others (8) it would seem that performance in this activity would be largely unique and only incidentally related

to basketball. The bounce and shoot test is scored according to the time taken to pick up a ball, dribble to the basket, shoot, recover and pass to a catcher on the other side of the court five times from each side of the court alternating from side to side. Separate scores are given for baskets made. As mentioned earlier it would seem doubtful if ten trials would be adequate to balance performance fluctuations on an accuracy test (8). The time scores might reflect certain aspects of basketball ability but have the disadvantage of not being diagnostic. It would be impossible to tell whether a low score was caused by lack of skill in dribbling, shooting, rebounding or running speed. A validity coefficient of .66 was found for the three items.

Clarke (6) reports the results of two batteries of basketball tests designed by Johnson. One test was a measure of basketball potential and the other was intended as a measure of basketball ability. The basketball potential test included: a footwork test, a jump and reach, a dodging run, and the Iowa Revision of the Brace Test (4). The basketball ability test consisted of: a thirty second field goal speed test, a dribbling agility run modified from the Edgren Basketball Test (11), and a basketball accuracy throw.

The principal difference between the Edgren dribbling test and the Johnson dribbling test is that the Edgren test is scored by the time taken to dribble through the agility run once while the Johnson test is scored in terms of the number of zones or obstacles passed in 30 seconds.

Johnson's basketball accuracy throw involves 10 baseball or hook passes from 40 feet at a series of rectangular targets 60 inches by 40 inches, 40 inches by 25 inches, and 20 inches by 10 inches. This

This item presents a number of problems. First it is unlikely that ten trials is sufficient to cancel out performance fluctuations especially in an accuracy test at this distance (8). Second, subjects are pre-mitted to use either the baseball or hook pass which presents an un-controlled factor. Third, neither the hook nor the baseball pass is among the most frequent passes used in basketball (1) and coaches prefer that most passes be thrown at a distance shorter than 40 feet (31). Fourth, with a rectangular target there is always the problem of logical validity because a subject could miss the target centre by the same distance at different angles and not receive the same score.

Johnson found the reliability and validity for the ability test were .89 and .88, respectively. But in securing validity he used the biserial correlation technique. Biserial correlation is a measure of the relationship between a continuous and a dichotomous variable (12). Johnson merely classified all of his subjects into two groups, good basketball players and poor basketball players. However, this form of analysis is not realistic when applied to basketball ability which is not a have or have not characteristic, but rather a wide continuum form a very poor to a very good.

Knox developed a basketball battery composed of a speed dribble, wall bounce, dribble and shoot, and the "penny-cup" test. The reliabil-ity coefficient for the battery as reported by Clarke (6) was .88 and for individual items ranged from .58 to .90.

Boyd, MCachren, and Waglow as reported in Clarke (6) found a bi-serial correlation of .96 between scores on the Knox Test and the

results of 42 candidates for the University of Florida junior varsity basketball squad with respect to whether they made the team or not. However, the correlations between test scores and the coach's ratings of the players were low.

Use of Statistics in Basketball. Wooden, (42) states that he finds great help in selecting his squad from keeping complete statistical charts and analysing them in a comparative manner. Statistical records used include a floor chart and tabulation of the number of field goals attempted and made plus percentage made, rebounds assists, personal fouls, and loss and gain of ball possession (42).

Case (5) believes that game and scrimmage statistics properly used can be a valuable aid to a coach. A player-team sheet comparing both teams and individuals helps to develop incentive and competition among players and between them and their individual opponents. Case (5) found from his experience that boys can be trained to do a fine job in using statistical and chart data. A record of shots attempted, number and percentage made, gained possession by recoveries or offensive and defensive rebounds, assists, lost possession, bad passes, fumbles, and offensive and defensive fouls presents a good summary of what a player has accomplished in a game.

Many coaches use statistics in evaluating their players but apparently, no one has systematized their statistics into a test form and researched their results to determine the validity, reliability and objectivity of their work.

<u>Statistical Analysis and Test Construction</u>. The criteria of a "good test" can be most easily understood by finding the answers to the following questions (6). Does the test measure the quality for which it is intended? Can the test be administered accurately? Can the test scores be interpreted in terms of relative performance? Is the test economical in terms of instrument cost and time involved?

According to Scott (37:241) the choice of test items by rational judgement is a more satisfactory method of test construction than selecting items through trial and error. "When constructing sports skill tests, the investigator should try to use items which are like the game situation." (37:242) Objective scoring systems are usually more reliable than subjective ones (37).

Validity is the degree to which the test measures the quality for which it is intended. Validity can be determined subjectively by what is termed curricular validty (37:224).

> A simple way to demonstrate curricular validity is to describe in detail the process of constructing the test. The author can offer proof of curricular validity at each of the several test construction steps.... the test will be valid to the extent that the items in the frame-work cover the subject of the test. If the literature shows that throwing, catching, running, and batting are important aspects of baseball, then the investigator demonstrates curricular validity when he lists these items in the framework and uses the literature to document their importance (37).

A teacher can show curricular validity for the tests he uses in his program by showing that the test items cover the course objectives (37). Another method of demonstrating curricular validity is by having a committee of judges evaluate the content of a test by considering: its functional values, importance of the

respective items, whether it produces the desired effect, suitability for the group to which it will be administered, whether it simulates a real situation, and the proper ratio of emphasis (37).

The proven validity of a test depends upon the correlation coefficient between test scores and a criterion (37). The Spearman correlation coefficient Rho is a measure of association between two variables which are expressed at least in an ordinal scale so that the objects or individuals under study may be ranked in two ordered series (39). When one has only a few scores (less than twenty-five) it is often advisable to rank the scores and use the Rho method of correlation instead of the more laborious product-moment r (16). Since Rho is an approximate measure it is hardly worth while computing a standard error of measurement (16).

A criterion is a known and accepted standard of whatever the examiner wishes to test. When established tests are known to be valid, they may be used as criteria to validate experimental tests (37). Competitive standings can be used as a criterion for a skill test(37). A test can be validated by using it on known groups which represent extremes in the quality to be measured (37).

The pooled ordering of judges' rankings may serve as a criterion for validation, especially when there is no relevant external criterion for ordering performance (39). Judges' ratings are probably the most common criterion for validating tests and are considered satisfactory if the judges are competent and have had an adequate opportunity to observe (37).

When we have k sets of rankings, we may determine the association among them by using the Kendall coefficient of concordance W.... Such a measure may be particularly useful in studies of inter-judge reliability and also has applications in studies of cluster variables (39).

The degree of agreement among the k judges is reflected by the degree of variance among the N sums of ranks. Therefore, the coefficient of concordance W is a function of that degree of variance (39). A high or significant value of W may be interpreted as meaning that the observers or judges are essentially in agreement in their rankings of the subjects under study. The Spearman Rho and the coefficient of coefficient of concordance W are linearly related. The average Rho for all possible combinations for which the coefficient of concordance W is calculated can be read from a table (39).

Sometimes, because no one method of validation contains all of the vactors desired as a criterion measure, an author will use a combination of methods of test validation (37).

Guilford (19:351) points out, "A coefficient of validity is always relative to the population sampled and to the manner in which the measurements were made." The size of the correlation coefficient is very much dependent upon the range of ability measured. The validity coefficient is almost invariably smaller in a restricted group than it would be in a relatively unrestricted group. Often we know the correlation between test scores and a criterion as derived from one group of individuals but shall be applying the same index to other groups with different ranges of ability. If one wishes to avoid wrong conclusions when a substantial amount of

selection has been made, one should apply correction procedures (19).

The correlation between a test and its criterion will be reduced if either the test scores or the criterion scores or both are unreliable (16). The correction for attenuation formula makes an adjustment for the effects of those chance or accidental errors which lower the reliability coefficients and thus affect the validity coefficient (16). Correction for chance errors will always raise the validity coefficient. However, if the reliability coefficient is lower than the validity coefficient, the correction for attenuation will yield a value greater than one. Such a result is logically meaningless and in such cases it is better to attempt to improve the reliability of the test than to correct for attenuation (19).

It is often impossible to get entire test validity. This would require an external criterion of achievement that is a better measure than the test itself and this is seldom available (38).

Often it is desirable to determine the statistical validity of individual test items. Item validity is a test to determine whether a given test item discriminates between known groups (16).

There are three dimensions to the statistical validity of test items; index of discrimination, difficulty rating, and functioning of responses (37).

The discriminating power of each item in the test is its ability to distinguish between students of varying abilities. The percentage passing a test item is its difficulty rating, the easier the item and vice versa. Functioning of responses refers to

the plausibility of alternate responses which does not apply to most motor performance tests.

A test item is said to discriminate when students who score high on it are found to achieve higher scores on the total test than students who score low on it (37). The experimenter discards those subjects in the middle of the distribution and works with the high and low scores (24). An item should be retained with an index of discrimination of twenty or higher if it meets other criteria (38).

A single test item has a perfect positive discriminating power if everyone in the high scoring group and no one in the low scoring group is successful; no discriminating power if there are equal proportions of success in the high and low groups; and negative discriminating power if there is a larger proportion of scores in the low group than in the high group (9).

There have been a number of methods of item validation advanced. The critical ratio between high and low groups estimates how certain we can be that the item discriminates. However, this method has the disadvantage of not allowing direct comparison of various items (9). It also is dependent on the number of subjects used (9).

Other methods utilize biserial, product-moment, or phi correlation coefficients. These methods are essentially unaffected by sample size but are drastically influenced by difficulty level (9). Items characterized by high coefficients are those which are passed successfully by fifty percent of the sample. These methods must be used with an index of difficulty to be meaningful

or the investigator will tend to use only items at the fifty percent difficulty level (9).

The Flanagan index of discrimination (14) is probably the bets known method of item validation. It yields a product-moment coefficient of correlation which indicates how well the item differentiates between good and poor performance. Because the Flanagan method uses a correlation coefficient, direct comparison of various items is difficult because correlation coefficients do not vary on a linear scale.

Davis has developed a table of discrimination indices based on Fisher's z transformation of the correlation coefficient multiplied by a constant (9). This provides indices for which the units are approximately equal throughout the scale and avoids the use of decimal points (7).

The Davis index of discrimination is truly comparable from item to item being independent of difficulty (9). A given increase has a constant meaning throughout the range (9). Since coefficients obtained by the procedure outlined by Kelley (24) are not affected by difficulty levels of the test items, and since essentially similar samples of testees are used, z values based on these coefficients may legitimately be added, subtracted and averaged.

Colver (7) has developed a nomograph from which both the Davis index of discrimination and the Flanagan index of discrimination can be determined from the percentage of successes in the high and low groups. Proportions of successes may be computed in several different ways. The use of the item analysis chart does not depend upon the employment of the recommended formula for computing proportions of success (9). Experience has shown that the use of nomographs result in saving a

significant amount of time with no loss in accuracy (7).

The magnitude of an item discrimination index is determined in part by the proportion of the total score that the item constitutes (9). If the total score is composed of items testing material similar to that in the given item, the discrimination index is apt to be high. If the item tests achievement in an area not closely related to that of the majority of the test, the index is apt to be low (9).

Reliability is the consistency of results obtained with a test administered by the same examiner. Reliability can be influenced by the time of day, the equipment, momentary attitude of the subject, lack of specific directions, and environmental conditions such as heat, light and humidity (37).

With the split half method of reliability the whole test is more reliable than either half and there is generally an increase in reliability with increased length of the test (19). The reliability of the whole test can be estimated by the Spearman-Brown formula (12). Guilford (19:493) asserts, "An estimate by the use of the Spearman-Brown formula is probably conservative, because it tends to be an underestimate."

If lengthening a test does not affect the range of difficulty and if boredom and fatigue effects are minimal, reliability can usually be improved. The Spearman-Brown prophecy formula can be used to estimate the effect of lengthening a test (16). The self correlation of a test is always affected by the variability of the group measured. The larger and more hetergeneous the group, the greater the test variability will be. The narrower the range of ability sampled, the smaller the correlation coefficient will be (19).

If we know the self correlation of the test in a narrow range we can estimate the self correlation of the same test in an increased range (16). Usually it is important to know the self correlation of a test in a larger population since it is probably more stabe from population to population than reliability coefficients from a restricted range (19).

Objectivity is the consistency with which a test can be administered by different examiners to the same subject or subjects (37). Objectivity is determined by having two examiners independently score the same group of subjects during the same test period. A correlation between the two sets of scores will yield an objectivity coefficient.

Averaging correlation coefficients is a dubious practice because correlation coefficients do not vary along a linear scale (16). The increase from r= .20 to r = .50 does not involve the same amount of increase as raising r = .80 to r = .90. Secondly, when positive correlations and negative correlations are averaged, they tend to cancel each other out. The best method of expressing a representative correlation coefficient for grouped data is to transform the individual correlation coefficients into Fisher's z functions and take the mean of the z's (13). The mean z is then converted into an equivalent correlation coefficient from a table (16).

If an examiner wishes to combine scores from a number of test items with different ranges and units of measurements he may convert the scores from each item into "T" scores which will make each item comparable (16). Then scores for each item will have the same meaning since reference is always to a standard scale of 100 units based on the normal probability curve and a standard deviation of 10 (16).

Summary of the Review of the Literature. The majority of basketball tests that have been proposed are based on opinions of the authors and are largely unsubstantiated by scientific evidence (6). Virtually all of the existing basketball tests are standardized skills tests which are not competitive. The research of Triplet, Berridge, Gordon and others as reported by Cratty (8) indicates that non-competitive and competitive performance is not comparable,

Most coaches would agree that execution of fundamentals in the game situation is what distinquishes between good and poor performers. Many coaches have developed elaborate techniques of keeping statistical records of the players' execution of fundamentals in games and scrimmages to aid their evaluation of the squad. However, no one has systematized their statistics into a test form and researched their results to determine the validity, reliability and objectivity of their system. There seems to be a real need for work in this area.

The criteria for a good test can be boiled down to four questions (6). Is the test valid? Is it reliable? Is it objective? And, can it be administered economically in terms of instrument cost and time involved?

CHAPTER III

METHODS AND PROCEDURES

Source of Data. Three sets of observations were made. The first set included twenty observations of intercollegiate varsity competition which consisted of three subsets. The first subset of observations was made during the regular 1965-66 Western Canada Intercollegiate Basketball Conference competition including all of the home games of the University of Alberta at Edmonton. These games provided six observations of the University of Alberta at Edmonton team and two observations each for the University of Alberta at Calgary, the University of Manitoba and the University of Saskatchewan. In this conference players are eligible to play for a maximum of five years while being enrolled as full time students. The teams play each other twice at home and twice away each year plus a varying number of exhibition games with other intercollegiate teams from both the United States and Canada.

The second subset of observations was taken from two controlled scrimmages involving the University of Alberta at Edmonton varsity team. The sample was a selected group and included the fourteen players carried on the team's roster. The first controlled scrimmage consisted of five periods of five minutes "stop time". The second controlled scrimmage consisted of six periods of five minutes stop time.

The third subset of observations was taken at the Fourth Annual Canadian Intercollegiate Basketball Championships held at the University of Alberta at Calgary on March 4th and 5th, 1966. The

data included two observations each for Acadia University, Carleton University, the University of Alberta at Calgary, and the University of Windsor. Each team was selected to represent the four respective zones of intercollegiate basketball in Canada on the basis of regular league standings.

The second set of data was taken from high school varsity games involving the following teams from the Edmonton, Alberta senior league: Bonny Doon Composite High School (one game). McNally Composite High (one game), and Strathcona Composite High School (three games). These teams were selected as an available sample. Players were full time students at the respective schools in grades ten, eleven, and twelve. Academic eligibility was at the discretion of the individual school.

The third set of data consisted of observations of forty=four members of two grade ten physical education classes from Strathcona Composite High School, Edmonton, Alberta. This included the entire population of the classes studied with the exception of those who were absent on the days of testing and an assistant in each group . Testing followed a five week unit on basketball in physical education class and for most students it was their second year of planned instruction in the activity. Ages of the subjects ranged from fourteen to sixteen. The classes studied were selected because of the availability of the sample and adequate facilities including a double basketball court. All observations involved male subjects.

Method of Obtaining the Data. The principal data gathering device used was the "Basketball Profile Statistics Chart" (see Appendix A). On the front of the profile sheet, there is a floor diagram bordered by boxes for the scoring of assists, freethrows

attempted and made, each way the ball can be gained, offensive and defensive rebounds, recoveries and positive held balls), and each way the ball can be lost (bad passes, fumbles, violations, negative held balls, and personal fouls).

For each shot attempted, the number of the player taking the shot was marked on the floor diagram, relative to the location from which the shot was taken. If the shot was made, the number was circled. Any time a player lost or gained the ball, the number of the player responsible (see definition of terms) was inserted in his team's portion of the appropraite box. Similarly, assists were recorded in the appropriate portion of the assist box. Positive and negative held balls were indicated by + and - signs. Foul shots attempted were entered below the floor diagram by the player's number and a circle for each shot taken. A single shot was indicated by a single circle, a double foul by two circles with an over bridge, and a one and one foul by a single circle with an under bridge plus a second circle if the first shot was made. Any time a foul shot was made the corresponding circle was filled in.

Four points for each filed goal made and two points for each free throw made, were awarded. One point was awarded for each rebound, positive held ball, recovery and assist. One point was deducted for each attempted field goal or free throw missed, bad pass, violation, fumble and personal foul.

Item totals were determined for each trait measured and were entered in the appropriate columns on the data sheets (see Appendix B). Item totals were converted to profile points by the regular (4,2,$\pm$1) point system described above and entered as subtotals under the headings:

field shots, free throws, rebounds, positive points, negative points and fouls. Total points for each player and team were tabulated and entered on the data sheets.

Experimental Design. Because there was little precedence to follow for this method of testing there did not seem to be a single criterion which contained all of the factors implicit in the profile device. Therefore a combination of methods of validation were used in the hope that each method might shed some light on the problem. Because of the unique nature of the test the principal method of validity analysis was through the curricular technique (37).

Three dimensions of curricular validity were demonstrated. First, it was shown that the material covered by the profile technique is important to successful performance in basketball competition. Second, it was shown that the various test items were scored with a logical ratio of emphasis. Third, it was shown that the Profile Test closely simulates a real situation.

Test results of the intercollegiate games were correlated with average rankings of a board of judges using Spearman's coefficient of rank correlation Rho (12). The board of experts for the Western Canada Intercollegiate Athletic Association Conference games included: four professors of physical education plus the head coach of the opponents of the University of Alberta at Edmonton. Two of the professors of physical education were former basketball coaches and one is currently coaching the varsity team at the University of Alberta at Edmonton. The judge who is currently coaching and the opposing coaches ranked only their respective opponents to avoid any bias they may have with their own teams.

The board of experts for the Canadian Intercollegiate Basketball Championships included: one professor of physical education who has previously been head coach of an intercollegiate basketball team, one physical education professor who is currently assistant coach of an intercollegiate team, one head coach of an intercollegiate team, and one judge who has played four years of intercollegiate basketball and who has represented Canada on the national basketball team.

The board of experts ranked players according to their overall performance in each game after observing the entire game.

The degree of agreement among the judges was determined by the coefficient of concordance W (39). If the coefficient of concordance was significant at the .05 level it was considered that there was sufficient agreement among the judges in ranking the players since the average rank of a board of experts could be expected to be more accurate than the ratings of a single judge. Validity coefficients were to be corrected for attenuation if the data justified the procedure according to the criteria outlined by Guilford (19).

The data from the Canadian Intercollegiate Basketball Championships was subjected to the same treatment as the data from the regular Western Intercollegiate Conference games.

Students from the high school classes were given the Modified Johnson Basketball Test (see Appendix C) for classification purposes. One practise trial and two trials to be scored were administered for each test item. Scores were coverted into standardized "T" scores so they could be combined (16). Students were matched with those nearest to them in ability according to the combined scores of the first and second trials for each test item.

The validity of the Modified Johnson Test was determined by correlating the results with the teacher's rankings using the Spearman correlation coefficient Rho (39) and corrected for attenuation as described by Garrett (16). The reliability of the Modified Johnson test was determined by correlating first trial scores with second trial scores and correcting for whole test length by means of the Spearman-Brown formula (19).

Profile scores were correlated with the teacher's rankings and with the results of the Modified Johnson Test by means of the Spearman Rho technique (39). The Kendall coefficient of concordance W (39) was also used to determine the degree of agreement among the Profile results, the Modified Johnson Test scores and the teacher's rankings.

Validity and reliability coefficients in the high school class study were depressed by the restriction of range caused by matching students with those nearest to them in ability from the results of the Modified Johnson Test. To determine the validity and reliability coefficients for the range of ability found in the classes sampled the correction procedures for restriction of range outline by Guilford (19) were used.

Each item's discriminating power was determined by the Davis technique (9) and the Flanagan technique (14) as read from the nomograph developed by Colver (7). Two point systems other than the regular (4, 2, $\pm$ 1) system were applied to the data to determine whether the discrimination of the various items and the test reliability could be improved. The first method was to award two points for rebounds, positive held balls, recoveries and assists, while deducting two points for negative held balls, bad passes, violations, fumbles, and fouls

while the point system for field goal and free throw shooting was the same as the regular point system. This point structure was called the 422 system.

The second method was to increase the number of points for each item which did not discriminate well according to the item analysis, while holding the number of points awarded to those items which discriminated well constant. The result was that the field goal, free throw, held ball, and foul points were unchanged. Two points were awarded for rebounds and recoveries, four points were awarded for assists and four points were deducted for bad passes, violations, and fumbles. This system was called the new point system.

The regular, 422, and new point systems were applied to both the intercollegiate results and the high school class situation.

Records of the time played for each player in five Western Canada Intercollegiate Athletic Association games were kept. Item scores from these games divided by the amount of time in seconds that each player performed were used for analysis of item discrimination for intercollegiate varsity basketball to minimize this variable factor.

Item analysis was made by comparing the best four players with the poorest four players on each team as determined by the total Profilo points divided by time played using each of the three point systesm applied. Comparisons between the results of the item analysis for the varsity teams were made with the results for the high school class study to determine if a single point system was adequate for both levels of basketball and to determine whether all items discriminated equally well at both levels of performance. An item was considered to discriminate

adequately if it yielded a Flanagan index of .20 or better as suggested by Scott (37,38).

Reliability for all profile scores was determined by correlating first half scores with second half scores by using the Spearman Rho (12) and estimating the reliability of the whole test by using the Spearman-Brown formula (19).

The reliability of the Profile technique as a measure of basketball "performance" (Cf ante p. 7) was determined for intercollegiate varsity basketball, high school varsity basketball and grade ten class basketball respectively by finding the split half reliability for each game. To determine the effects of the amount of time played on performance. Profile scores were divided by the amount of time played by each player for five Western Canada Intercollegiate Athletic Association games. Reliability coefficients for the high school class study were corrected for restriction of range as suggested by Guilford (19).

The reliability of the Profile technique as a measure of basketball "ability" (Cf ante p. 5) for intercollegiate varsity basketball was determined by finding the reliability over a number of games. All combinations of two, three, four, five and six game profile scores for the University of Alberta at Edmonton varsity team were correlated first halves vs second halves and corrected by the Spearman-Brown Formula (12).

Objectivity was determined by correlating the author's test scores for the Western Intercollegiate Conference Games with the test scores obtained independently by the managers of the University at Edmonton varsity team. A second objectivity study was conducted by

correlating the author's test scores for the high school varsity games with the scores obtained independently by the manager of the Strathcona senior high school varsity team. The Spearman Rho correlation technique was used (12).

The amount of training necessary for an observer to obtain accurate results was determined by the number of games an observer must score to obtain an objectivity coefficient significant at the .01 level.

Two time studies were undertaken in order to determine if the duration of test necessary to obtain a corrected split half reliability of .85 or better is practical with due respect for fatigue and administrative considerations (37). Two controlled scrimmages involving the University of Alberta at Edmonton varsity team were used.

The first controlled scrimmage consisted of five periods of five minutes "stop time" each. The second controlled scrimmage consisted of six periods of five minutes stop time each. The various periods were grouped into sets of five, ten and fifteen minute periods. These sets were correlated with all possible mutually exclusive periods of the same length. The Spearman-Brown prophecy formula was used to estimate the effect of lengthening the test (19).

Because so many correlation coefficients were determined it was necessary to group data by means of the Fisher z transformation (13). The probability values associated with the resulting coefficients were determined by the appropriate tests of significance outlined by Siegel (39).

The decision concerning the suitability of the test for the evaluation of: intercollegiate varsity basketball, high school varsity

basketball and high school class basketball was determined by whether the achieved probability values were significant at the .01 level. Similarly, the suitability of the test as a grading device was determined by the probability values determined from the reliability and validity measure taken in the high school class study.

All formulae used in the study may be found in Appendix G.

CHAPTER IV

RESULTS AND DISCUSSION

<u>Curricular</u> <u>Validity</u>. The Profile technique embodies curricular validity in three dimensions. First, the material covered by the Profile technique is important to successful performance in basketball competition. Each test item embodies successful or unsuccessful execution of basketball fundamentals. Dean (10), Holman (22), McCracken (31) and others have stated that basketball is a game which is dependent upon successful execution of fundamentals.

Mauer as cited in Dean (10) has stated that possession of the ball means control of the game. Offensive and defensive rebounds, positive held balls, and recoveries are test items which reward gaining possession of the ball. Deductions for missed field goals, missed free throws, negative held balls, bad passes, violations, fumbles, and fouls penalize losing possession of the ball.

Dean (10) argues that the only substitute for good ball handling is superior shooting ability. The test of an offense is its ability to control the ball while working into position for a good percentage shot. Dean (10) emphasized that it has been estimated that about fifty percent of all basketball games are won or lost at the free throw line. The Profile test has items covering both field shooting and free throw shooting.

McCracken (31) feels that passing, shooting and gaining possession of the ball through held balls, rebounds and recovery of "free balls" is so essential to basketball performance that he devoted separate chapters to each of these fundamentals. As mentioned above,

both field shooting and free throw shooting are covered by test items. Passing is covered by both a negative and a positive scale. Bad passes are penalized by point deductions while good passes which enable a teammate to take an unhindered shot in good scoring range are awarded points under the assist heading. Each method of gaining possession of the ball discussed by McCracken (31) is covered by test items.

McCracken (31) warns against losing possession through poor shots, inaccurate passes, and being forced into negative held ball situations. Each of these mistakes are penalized by negatively scored test items.

Dean (10) strongly emphasizes that most fouls are careless and can be prevented. Coaches should teach their players how to avoid fouls. The Profile technique discourages fouling by penalizing the player for each foul committed.

Basketball experts (1,10,31 and 42) list a number of other factors which help to make a successful basketball player. Most of these factors involve good position playing such as: good defensive position, blocking out on rebounds, and screening. These factors are difficult to quantify objectively but it is considered that good position will lead to successful performance of Profile items. Good screens will be awarded as assists. Good "blocking out" will result in more rebounds. Good defensive position will result in more recoveries and positive held balls. Good offensive maneuvers will give a player more good shots. For these reasons it is considered that the Profile technique, measures the essential fundamentals of basketball.

The second dimension of the Profile technique's curricular validity is the logical ratio of the point system for the various items. Mauer as cited in Dean (10) and McCracken (31) have indicated, possession

of the ball means control of the game. As McCracken (31) emphasized, gaining possession of the ball is essential to successful basketball performance. It would seem logical that each method of gaining the ball should be scored alike since the resulting benefit to the team is the same (i.e. possession of the ball). Therefore offensive and defensive rebounds, positive held balls, and recoveries were all awarded one point.

It seemed reasonable that any mistake that causes loss of possession should be penalized equally. Therefore, missed field shots, missed free throws, negative held balls, bad passes, violations, fumbles and fouls all resulted in the same point deduction. Since losing possession placed the team in the same position as they were before possession was gained, it was logical that the deduction for losing possession of the ball should be the same as the number of points awarded for gaining possession. Each way in which possession could be lost was penalized by a one point deduction.

As Dean (10) suggests, the only fundamental that could be regarded as more important than ball-handling is good shooting. McCracken (31) is even more emphatic. He argues that when one thinks of a basketball player, there are many things to take into consideration, but whether the player is a scoring threat or not is the ultimate question.

Good shooting is the only fundamental which was awarded more points than ball handling. Dean (10) emphasizes that the test of an offense is its ability to control the ball while attempting to advance the ball into good percentage shooting position.

The Profile point system is based on the percentage of shots scored. The greater the ratio of net points scored and vice versa. A missed shot resulted in lost possession of the ball just as much as a bad pass, violation, fumble or foul unless the shooter rebounded his own shot. Therefore, a missed shot was penalized by one point just as other methods of lost possession were penalized. If the shooter rebounded his own shot the penalty for missing the shot was cancelled our by the one point awarded for the rebound. Because four points are awarded for each shot scored and one point is deducted for each shot missed the net points per basket made varied with the percentage of shots made. Table I below demonstrates how the net values awarded for selected shooting percentages varied.

TABLE I

NET POINTS PER FIELD GOAL

Percentage Made	Net Points Per Field Goal	Percentage Made	Net Points Per Field Goal
100	4.00	50	3.00
80	3.75	33	2.00
75	3.67	25	1.00
67	3.50	20	0.00
60	3.33	10	- 4.00

Because free throws are awarded half as many points as field goals in basketball it was logical that for each free throw made, half as many Profile points should be awarded for free throws as for field goals, However, each time a free throw was missed, possession was lost just as when a field shot was missed. Therefore, one point was

deducted for a missed free throw. If the shooter rebounded his own shot the point awarded for the rebound cancelled out the one point deduction off the shot missed because possession of the ball was not lost. As with field shots, net points awarded for free throws made varied with the percentage of shots made. In table II below, a number of values awarded for various shooting percentages is listed.

TABLE II

NET POINTS PER FREE THROW

Percentage Made	Net Points Per Free Throw	Percentage Made	Net Points Per Free Throw
100	2.00	50	1.00
80	1.75	33	0.00
75	1.65	25	- 1.00
67	1.50	20	- 2.00
60	1.33	10	- 7.00

The third dimension of the Profile technique's curricular validity was the realistic situation that was embodied in the Profile method of testing. As a measure of basketball performance, all overt acts that could be objectively scored in a quantitative form from the actual game situation were used in the evaluation. As a test, the Profile technique was used in a situation that was exactly the same as a game situation except that one's opponents and teammates were determined by the examiner and time was controlled so that each player played the same length of time.

Proven Validity. The Profile results for the intercollegiate games were correlated with the average rankings of a board of experts. A

summary of the average validity coefficients for the intercollegiate games follows in table III. Individual coefficients were converted to Fisher's z transformations (13), added, averaged and reconverted to validity coefficients in the manner advised by Garrett (16). A detailed list of all validity coefficients may be found in Appendix D.

From table III it can be seen that the average Rho for all intercollegiate games was .785. This result was determined from twenty validity coefficients with a median N of 9 and was significant at the .01 probability level. The average Rho for the Western Canada Intercollegiate Athletic Association games studied was .770. This quantity was derived from twelve Rhos with a median N of 9 was not significant at the critical value of .01. However, it should be noted that this coefficient was within .013 of being significant at the .01 level. For the Canadian Intercollegiate Athletic Association Championship games the average Rho was .805. Eight Rhos were averaged to give this value with a median N of 9. This correlation was significant at the .01 level of confidence.

Average Rhos were computed for the individual teams studied and are reported in table III.

TABLE III

SUMMARY OF INTERCOLLEGIATE VALIDITY COEFFICIENTS

Study	Team	Number of Rhos	N	Rho	Critical Values
Combined	All teams	20	9	.785	.01
* W.C.I.A.A.	All teams	12	9	.770	.05
**C.I.A.A.	All teams	8	9	.805	.01
W.C.I.A.A.	University of Alta. (Edmonton)	6	9	.785	.01
W.C.I.A.A.	Opponents of U. of A. (Edmonton)	6	8	.755	.05
W.C.I.A.A.	University of Saskatchewan	2	7	.735	.05
W.C.I.A.A.	University of Manitoba	2	10	.770	.01
W.C.I.A.A.	University of Alta. (Calgary)	2	8	.760	.05
C.I.A.A.	University of Alta. (Calgary)	2	10	.850	.01
Combined	University of Alta. (Calgary)	4	9	.810	.01
C.I.A.A.	University of Windsor	2	10	.895	.01
C.I.A.A.	Acadia University	2	8	.850	.01
C.I.A.A.	Carleton University	2	7	.425	not significant

* W.C.I.A.A. - Western Canada Intercollegiate Athletic Association

** C.I.A.A. - Canadian Intercollegiate Athletic Association

It should be noted that the average Rho for the University of Alberta at Edmonton was .785 and significant at the .01 level but the average Rho for the opponents of the University of Alberta was .755 which is not significant at the .01 but the .05 level. Also it should be noted that the mean Rho for the University of Alberta at Calgary for the Western Canada Intercollegiate Conference games was .760 which was not significant at the .01 level, but for the Canadian Intercollegiate Championships the mean Rho for the University of Alberta at Calgary was .850 which is significant at the .01 level of confidence.

The degree of agreement among the judges was determined by the coefficient of concordance W as described by Siegel (39). The significance of W was determined by a Chi square test as described by Ferguson (12). The average Rho for all possible combinations for which the coefficient of concordance W was calculated was obtained from a table (16:154). In table IV the coefficients of concordance W for all intercollegiate games studied, the critical value of each W, the Chi square value, and the equivalent mean Rho are listed.

It can be seen that eleven of the nineteen W's calculated were significant at the desired probability level ($p = .05$) and that eight were not significant. The median level of significance was at the desired probability level ($p = .05$). It should be noted that the median critical value for the home team, the University of Alberta, at Edmonton, was .02 whereas the median values for the opponents was .20.

TABLE IV

SIGNIFICANCE OF THE COEFFICIENT OF CONCORDANCE W

Home Team	W	N	Rho	Chi^2	P Value	Opponent	W	N	Rho	Chi^2	P Value
U. of A.	.645	11	.527	25.8	.01	U. of S.	.292	7	.115	8.76	.20
U. of A.	.831	8	.747	17.45	.02	U. of S.	.254	7	.005	6.10	.50
U. of A.	.946	9	.892	15.14	.10	U. of M.	.941	10	.882	16.94	.05
U. of A.	.959	9	.939	23.02	.01	U. of M.	.696	7	.392	8.35	.30
U. of A.	.757	9	.636	18.17	.02	Calgary	.816	8	.724	17.14	.02
U. of A.	(insufficient Judges)					Calgary	.714	7	.428	8.57	.20
Windsor	.679	10	.519	18.33	.05	Acadia	.958	7	.937	17.24	.01
Carleton	.892	8	.838	18.73	.01	Calgary	.770	9	.655	16.27	.05
Carleton	.857	6	.714	8.57	.20	Acadia	.750	9	.500	12.00	.20
Windsor	.955	10	.910	17.19	.05	Calgary	.832	10	.664	14.98	.10

From observation of the data it would appear that the Profile technique was more sensitive to performance fluctuations than were the board of experts. To test this hypothesis, the range of ranks for each player on the University of Alberta (Edmonton) team were determined for both the experts' rating and Profile scores. The mean range of ranks as determined by the Profile scores was 7.56 while the mean range of ranks as determined by the experts' ratings was 5.28. A t-test for determining the differences between two means for a correlated sample (12) indicated a significant difference ($p = .05$).

Students from high school classes studied were given the Modified Johnson Basketball Test (see Appendix C) for classification purposes. The reliability of the Modified Johnson Test as administered to the two classes was .894 and .911 with an N of 20 and 24 respectively. Both correlations were significant at the .01 level. The validity coefficients for the two classes as determined by correlating the results of the Modified Johnson Test with the teacher's rankings were .836 and .815 with N equal to 18 and 20 respectively. Both validity coefficients were significant at the .01 level of confidence.

Class Profile scores were correlated with the teacher's rankings and with the results of the Modified Johnson Test. The average validity coefficient as determine by correlating the Profile rankings with the teacher's rankings was .530 in the highly restricted range where each subject was matched with the classmate closest to him in ability and placed in a group with teammates and opponents similarly chosen. The mean validity coefficient for the

larger range of ability found in the whole class was .884 with N = 20 (significant with P = .01).

The mean validity coefficient as determined by correlating the Profile scores with the results of the Modified Johnson Test was .475 for the highly restricted range but in the larger range of ability found in the whole class the validity coefficient was .870 with N = 20 (significant at the .01 level). Detailed results of the validity analysis of the high school class study may be found in Appendix E.

The degree of agreement among the Profile results, the Modified Johnson Test scores and the teacher's rankings was found by the coefficient of concordance W (39).

Table V illustrates the values of W obtained, the equivalent mean Rho, the Chi square value associated with the obtained value of W and the level of significance. It can be seen that although three out of the eight groups have values significant at the .05 level indicating agreement among the three methods of evaluating basketball performance, five of the eight values were not significant at the .05 level indicating that there was no essential agreement among the three methods of evaluating basketball performance.

TABLE V

SIGNIFICANCE AND EQUIVALENT RHO'S OF COEFFICIENT OF CONCORDANCE FOR THE PROFILE TEST, MODIFIED JOHNSON TEST AND TEACHER'S RANKINGS FOR GRADE TEN HIGH SCHOOL BOYS

Class	Group	Team	W	Rho	Chi^2	Significance
3 and 4	I	Green	.306	-.041	3.67	.50
3 and 4	I	Gold	.778	.667	9.34	.10
3 and 4	II	Green	.689	.534	8.27	.10
3 and 4	II	Gold	.911	.867	10.93	.05
5 and 6	I	Green	.889	.834	10.67	.05
5 and 6	I	Gold	.822	.733	9.86	.05
5 and 6	II	Green	.644	.466	7.73	.20
5 and 6	II	Gold	.083	-.376	1.00	.95

<u>Item Analysis</u>. The Davis Index (9) and the Flanagan Index (14) were used to determine the ability of each test item to discriminate between good and poor performers. Two point systems other than the regular (4, 2, ± 1) system were applied to the data to determine whether the discrimination of the various items could be improved. The first method was called the 422 system (Cf. ante p.29).

The second method was called the new point system (Cf. ante p. 30).

The regular, 422, and new point systems were applied to both the intercollegiate results and the high school class situation. Table VI portrays the discrimination values in both the Davis and Flanagan scales for each Profile item using the regular, 422, and new point systems for the intercollegiate item scores divided by time played.

From table VI it can be seen that the field shot item discriminated very well with all point systems but discriminated best with the regular point system. The free throw item discriminated adequately in all three point systems but discriminated best with the 422 system. The offensive rebound item discriminated well with regular system, adequately with the new point system but yields a negative index with the 422 point system. The regular point system is slightly superior to the new system in yielding discriminating values for defensive rebounds but the 422 point system did not discriminate adequately on this item. Held balls was a discriminating item with the regular point system but failed

to discriminate adequately with the 422 or new point systems.

Recoveries did not yield an index of discrimination of .20 or better with any of the three point systems applied. Because the difference between recoveries and rebounds is largely a matter of definition it was decided to determine whether offensive rebounds, defensive rebounds, recoveries and positive held balls as a cluster variable designated as take-overs discriminated adequately. It was found that take-overs yielded a Davis index of 18 and an equivalent Flanagan index of .30. This is to be compared to a Davis index of 15 and an equivalent Flanagan index of 24 as determined by averaging the Davis indices for offensive and defensive rebounds, positive held balls and recoveries using the regular point system.

Assists yelded a high index of discrimination with the 422 point system but failed to discriminate adequately with the regular or new point systems. The index of discrimination for bad passes was barely adequate with the 422 point system but did not discriminate with either the regular or new point systems. Violations yielded the best power of discrimination with the new point system and inadequate discrimination with the regular point system.

Fumbles did not yield an adequate index of discrimination with any of the three point systems applied. The 422 system yielded the best negative values however and the regular point system showed positive discrimination for this negative item. Fouls yielded negative discriminating values in all three point ratios but the 422 system yielded the best discrimination with this item.

Each of the three point systems demonstrated adequate discriminating power on six items. The regular point system provided

the best discrimination on five items while the 422 system was the best selector on four items. The new point system was the best selector on two items. It should be noted that the regular point system consistently showed better discrimination on positive items while the 422 point system was consistently superior with the negative items.

Table VII illustrates the discrimination indices in both the Davis and Flanagan scales for each Profile item using the regular, 422 and new point systems for the high school class study.

TABLE VI

PROFILE ITEM DISCRIMINATION
FOR INTERCOLLEGIATE VARSITY BASKETBALL

Test Items	New Point System Davis	New Point System Flanagan	Regular Point System Davis	Regular Point System Flanagan	422 Point System Davis	422 Point System Flanagan
Field Shots	53	.71	67	.82	64	.79
Free Throws	18	.30	16	.27	21	.34
Offensive Rebounds	14	.23	20	.33	- 12	- .20
Defensive Rebounds	15	.24	16	.27	4	.07
Held Balls	- 1	- .01	17	.28	8	.13
Recoveries	1	.01	7	.11	0	.00
Assists	- 1	- .01	7	.11	23	.38
Bad Passes	-10	- .16	- 1	- .01	- 12	- .20
Violations	-21	- .34	- 10	- .16	- 17	- .27
Fumbles	- 2	- .02	16	.27	- 4	- .06
Fouls	-21	- .34	- 12	- .23	- 24	- .38

NB. underscored figures indicate adequate discriminating power.

TABLE VII

PROFILE ITEM DISCRIMINATION FOR
HIGH SCHOOL GRADE TEN CLASS BASKETBALL

Test Items	New Point System Davis	New Point System Flanagan	Regular Point System Davis	Regular Point System Flanagan	422 Point System Davis	422 Point System Flanagan
Field Shots	100	.93	100	.93	100	.93
Free Throws	100	.93	100	.93	100	.93
Offensive Rebounds	30	.46	30	.46	35	.50
Defensive Rebounds	16	.26	25	.41	23	.38
Held Balls	17	.27	44	.64	26	.43
Recoveries	20	.32	16	.27	22	.36
Assists	0	.00	- 13	- .31	- 5	.08
Bad Passes	- 9	- .15	- 11	- .18	- 5	- .08
Violations	- 14	- .22	15	.25	0	.00
Fumbles	- 8	- .13	- 6	- .10	- 8	- .13
Fouls	38	.56	39	.58	51	.69

NB. underscored figures indicate adequate discriminating power.

All three point systems yielded perfect discrimination with the field shot and free throw items. In each case the high group had positive scores and the low group had negative scores. Therefore it is necessary to look at the raw data to determine which system discriminates best on these items.

TABLE VIII

PROFILE SCORES FOR HIGH AND LOW GROUPS
OF GRADE TEN BOYS ON THE FIELD SHOT AND FREE THROW ITEMS

Point System	Field Shots High Group	Field Shots Low Group	Free Throws High Group	Free Throws Low Group
Regular Point System	72	- 51	6	- 7
422 Point System	62	- 37	7	- 4
New Point System	60	- 35	5	- 1

It can be seen from table VIII that the regular point system yields a greater point spread between high and low groups for both field shots and free throws than either the 422 or the new point systems.

Offensive rebounds discriminate very well in all three point systems but discriminate best with the 422 point ratio. Defensive rebounds discriminate adequately in all three point scales but discriminate best with the regular point system. Held balls also discriminate adequately with all three point systems but discriminate much better with the regular ratio than with the other systems. Recoveries discriminate with all three point systems. The new point system discriminates best on this item.

Assists do not discriminate with any of the three point systems and yield negative values with the regular and 422 systems. It appears the new point system is the best with this item and it does not discriminate.

Bad passes did not discriminate adequately with any of the three point systems. The regular point system came closest to adequate discriminating power with a negative value of .18. Violations discriminated adequately with the new point system but did not discriminate,

with the 422 system and yielded a positive discrimination for the negative item with the regular point system. All three point systems produced inadequate discrimination for fumbles. The 422 and new point systems were identical with Flanagan indices of -.13 while the regular system was slightly less discriminating with an index of -.10.

With each point system applied, the fouls item produced large positive indices for this negative item. The new and regular point systems were very close with Flanagan indices of .56 and .58 respectively and the 422 point system yielded an index of .69.

In the high school class study, the new point system yielded discriminating scores on seven items while the 422 and regular point systems discriminated on six items each. The regular point system was the best discriminator on five items and the 422 and new point systems were the best discriminators on three items and two items respectively. One item was equally discriminating with either the 422 or the new point system. Only one of the items for which the new point system was the best selector discriminated at the criterion level of .20 or better on the Flanagan scale while four of the items for which the regular system was the best selector discriminated at the criterion level of .20 or better.

Reliability. The reliability of the Profile technique as a measure of basketball performance was determined using the Spearman Rho technique (12) to correlate first and second halves. The average reliability coefficients as determined by Fisher's z transformation (13) for all intercollegiate varsity observations are portrayed in table IX. The average reliability coefficient for all twenty observations was

found to be .596 with a median N of 8. The average reliability coefficient for the Western Canada Intercollegiate Athletic Association games studied was .577 with a median N of 8. The Canadian Intercollegiate Athletic Association Championships yielded and average reliability of .621 with N = 9; University of Saskatchewan, .440 with N = 7; University of Manitoba, .310 with N = 8; University of Alberta at Calgary, .621 with N = 9; University of Windsor, .690 with N = 9; Acadia University, .861 with N = 6; and Carleton University, -.100 with N = 7. None of the reliability coefficients were significant at the critical value of .01. Reliability coefficients for each observation are listed in Appendix F.

TABLE IX

RELIABILITY COEFFICIENTS FOR THE PROFILE TECHNIQUE
AS A MEASURE OF BASKETBALL PERFORMANCE
FOR INTERCOLLEGIATE VARSITY BASKETBALL

Sample	Reliability Coefficient	N Subjects	N Observations	Critical Value
Total	.596	8	20	.833
W.C.I.A.A.	.577	8	12	.833
C.I.A.A.	.621	8	8	.833
U. of Alberta	.630	9	6	.783
U. of Saskatchewan	.440	7	2	.893
U. of Manitoba	.310	8	2	.833
U. of A. Calgary	.621	9	4	.783
U. of Windsor	.690	9	2	.783
Acadia U.	.861	6	2	.943
Carleton U.	-.100	7	2	.893

The Profile scores were divided by the amount of time played by each subject for the five Western Canada Intercollegiate Athletic Association games which were timed. Table X presents the reliability coefficients for the five games timed with and without scores being divided by time. The average reliability coefficient for scores being divided by time was .430 while the average reliability coefficient for scores not divided by time was .555.

It should be noted that the average reliability coefficients for the University of Alberta (.670 with time partialled out and .695 with time not partialled out) are considerably more reliable than the average Spearman Rho coefficients for the opponents (.105 with time partialled out and .460 with time not partialled out).

The sum of squares of the differences between paired ranks for the upper half (regulars) and the lower half (substitutes) of the subjects according to time played was calculated to determine whether comparatively small errors for those who played for short periods of time were being magnified when scores were equated for time played. From table XI it can be seen that the sum of squares of the differences between paired ranks for the upper half (regulars) and lower half (substitutes) were identical on one occassion, higher for the upper half on four occassions and higher for the lower half on five occassions.

TABLE X

RELIABILITY COEFFICIENTS WITH AND WITHOUT TIME PARTIALLED OUT FOR FIVE INTERCOLLEGIATE VARSITY GAMES

Home Team	Time Partialled Out	Time Not Partialled Out	Opponent	Time Partialled Out	Time Not Partialled Out
U. of Alberta	.778	.730	U. of Sask.	-.250	.310
U. of Alberta	.746	.610	U. of Manitoba	-.486	-.700
U. of Alberta	.700	.520	U. of Manitoba	.698	.670
U. of Alberta	.168	.730	U. of A. Calgary	-.132	.490
U. of Alberta	.759	.820	U. of A. Calgary	.526	.830
Average (U. of A.)	.670	.695	Average (Opponents)	.105	.460
All Observations	.430	.555			

TABLE XI

THE SUM OF THE DIFFERENCES SQUARED
OF REGULARS AND SUBSTITUTES
BETWEEN FIRST AND SECOND HALF PROFILE SCORES DIVIDED BY TIME

Team	Σd^2 Regulars	Σd^2 Substitutes
U. of Alberta	29	1.5
U. of Alberta	29	.5
U. of Alberta	29	15
U. of Alberta	39	69
U. of Alberta	9	9
U. of Saskatchewan	20	20
U. of Manitoba	29	44
U. of Manitoba	13	26
U. of A. Calgary	54	36
U. of A. Calgary	6	21

The reliability of the Profile technique as a measure of performance for high school varsity basketball was determined by the split half method of correlation with the Spearman coefficient of rank correlation Rho. The average reliability coefficients as determined by Fisher's z transformation (13) for high school varsity games are listed in table XII.

TABLE XII

RELIABILITY COEFFICIENTS FOR THE PROFILE TECHNIQUE
AS A MEASURE OF BASKETBALL PERFORMANCE
FOR HIGH SCHOOL VARSITY BASKETBALL

Team	Reliability Coefficient	N Subjects	N Observations	Critical Value (.01)
Strathcona	.66	7	3	.893
McNally	.41	6	1	.943
Bonny Doon	.48	8	1	.833
Average	.59	7	5	.893

The average reliability for the total group was .59. Composite High School's reliability coefficient was .66. McNally Composite High School and Bonny Doon Composite High School had reliabilities of .41 and .48 respectively. None of the average reliabilities were significant. However for one game, Strathcona had a reliability coefficient of .95 which was significant at the .01 level.

The reliability of the Profile technique as a measure of basketball performance for grade ten high school physical education for a twenty minute period of "straight time" was calculated for both the restricted range where each subject was matched with the classmate closest to him in ability and placed in a group with teammates and opponents similarly chosed, and for the larger range of ability found in the whole class. Table XIII presents the average reliability coefficients for both the restricted range and the entire class for the regular, new, and 422 point systems (Cf. ante p.30) as determined by Fisher's transformation (13).

TABLE XIII

RELIABILITY OF THE PROFILE TECHNIQUE AS A MEASURE OF BASKETBALL PERFORMANCE FOR GRADE TEN HIGH SCHOOL BOYS

Range	Regular Point System	New Point System	422 Point System
Restricted	.410	.360	.470
Class	.854	.842	.869

The reliability for the high school class study in the highly restricted range using the regular point system was .410. In the larger range of the whole class, the reliability was .854 with N = 20 (significant at the .01 level). With the new point system the reliability coefficient was .360 in the narrow range and .842 for the whole class (significant at the .01 level). The 422 point system yielded coefficients of .470 and .869 (significant at the .01 level) for the restricted range and the whole class respectively. The reliability coefficients as determined by the three point systems were not significantly different.

The reliability of the Profile technique as a measure of basketball "ability" (Cf. ante p. 5) for intercollegiate varsity basketball was determined by finding the reliability for the first halves vs the second halves for all combinations of two, three, four, five and six game Profile scores for the University of Alberta at Edmonton varsity team.

From table XIV it can be seen that the reliability over a period of six games is .838. Over a period of five games the reliability is .795 and for a four game period the reliability is .750. These three

values are significant at the .01 level. However, the reliability over a period of three games and two games was .720 and .695 respectively. Neither value is significant at the .01 level. The average reliability for single games for the University of Alberta was .630.

TABLE XIV

RELIABILITY OF THE PROFILE TECHNIQUE
AS A MEASURE OF BASKETBALL "ABILITY"
FOR INTERCOLLEGIATE VARSITY BASKETBALL

Number of Games	Realiability Coefficient	Significance
6	.838	.01
5	.795	.01
4	.750	.01
3	.720	.05
2	.695	.05
1	.630	.05

Objectivity. Objectivity was determined by correlating the author's Profile scores with scores obtained independently by student managers. The Western Canada Intercollegiate Athletic Association games were scored by the student manager for the University of Alberta at Edmonton. Table XV below presents the resulting objectivity coefficients. The average coefficient for the University of Alberta at Edmonton and opponents combined was .910 with N = 9 (significant at the .001 level).

The average objectivity coefficient for the University of Alberta at Edmonton was .950 with N = 10 (significant at the .001 level)

while the coefficient for the opponents was .875 with N = 9 (significant at the .01 level). It should be noticed that the average objectivity coefficient for the home team was larger and significant at a higher critical value than was the average coefficient for the opponents.

TABLE XV

OBJECTIVITY COEFFICIENTS BETWEEN AUTHOR"S AND STUDENT MANAGER"S PROFILE SCORES FOR INTERCOLLEGIATE VARSITY BASKETBALL

Team	Objectivity Coefficient	N	Average Coefficient	Significance
All teams		9	.920	.001
U. of Alberta	.911	11		.001
U. of Alberta	.967	9		.001
U. of Alberta	.954	9		.001
U. of Alberta	.936	10		.001
U. of Alberta	.967	10		.001
U. of Alberta		10	.950	.001
Opponents of U. of Alberta		9	.875	.01
U. of Saskatchewan	.938	7		.01
U. of Saskatchewan	.857	7		.02
U. of Saskatchewan		7	.905	.01
U. of Manitoba	.800	10		.01
U. of Calgary	.863	9		.01
U. of Calgary	.883	9		.01
U. of A. Calgary		9	.875	.01

High school varsity games were scored independently by the author and the student manager of Strathcona Composite High School. From table XVI it can be seen that the average coefficient for all teams for three games was .910 with N = 9 (significant at the .001 level). Results for Strathcona Composite High School, the home team, were .930 with N = 9 (significant at the .001 level) while the average coefficient for the opponents was .880 (significant at the .01 level).

TABLE XVI

OBJECTIVITY COEFFICIENTS BETWEEN THE AUTHOR"S AND STUDENT MANAGER'S PROFILE SCORES FOR HIGH SCHOOL VARSITY BASKETBALL

Team	Objectivity Coefficient	N	Significance
Strathcona	.930	9	.001
McNally	.780	8	.05
Bonny Doon	.950	11	.001
All teams	.910	9	.01

Training Observers. The amount of time required to train an observer to obtain accurate results was determined by the number of games an observer must score to obtain an objectivity coefficient significant at the .01 level. Table XVII illustrates the number games required to obtain objective results by both a university manager and a high school manager.

TABLE XVII

THE NUMBER OF GAMES AN OBSERVER MUST
SCORE TO OBTAIN OBJECTIVE RESULTS

Group	University Manager	High School Manager
Home Team	1 game	2 games
Unfamiliar Team	3 games	3 games

It should be noted that both high school and university students can use the Profile technique with accuracy at the critical level of .01 on the third exposure. The average objectivity coefficient for five games as scored by the university manager was .920 with N = 9 (significant at the .001 level) including the first two games.

The high school manager achieved an average objectivity coefficient for the home team of .930 (significant at the .001 level) and an average objectivity coefficient for the opponents .880 (significant at the .01 level) for the first three games that he scored.

Time Studies. Two time studies were made to determine the effect of the time variable. A summary of the results of these studies may be found in table XVIII below. It can be seen that the reliabilities for ten minute periods were .174 and .238 for the first and second time studies respectively. The projected reliability for twenty, thirty and forty minute periods as estimated by the Spearman-Brown prophecy formula (19) were: .337, .438, and .509 respectively. In time study I the achieved reliability was .312 which was reasonably close to the estimate by the Spearman-Brown

formula but the achieved reliability in time study II was -.086 for twenty minutes and -.671 for thirty minutes, in direct opposition to the predictions.

TABLE XVIII

SUMMARY OF RESULTS FOR TWO TIME STUDIES
FROM CONTROLLED SCRIMMAGES
WITH THE UNIVERSITY OF ALBERTA VARSITY TEAM

Time	Study I	Study II	Spearman-Brown Estimate
10	.174	.238	
20	.312	-.086	.337
30		-.671	.438
40			.509

Discussion

Curricular Validity. Three dimensions of curricular validity were shown for the Profile technique of evaluating basketball performance. First, items were selected by rational judgement as recommended by Scott (37) to measure performance in fundamental skills which have been designated as essential to successful participation in the game of basketball by recognized authorities on the subject. Second, the point ratio was shown to reflect the functional value and relative importance of the various items. Third, it was shown that a game like situation was not only implicit in the Profile method of testing but specific game situations could be imposed within the testing procedure by instructing the players to use systems of offense or defense learned in class or practice.

The research of Triplet, Berridge and Gordon as reported by Clarke (8) would imply that basketball performance in a series of non-competitive isolated drills would be quite different from performance in a competitive, integrated game situation. Unlike the standardized skill tests which measure a series of isolated skills in a non-competitive situation, the Profile Test measured competitive performance in a game situation.

Proven Validity. Proven validity for the Profile technique as a measure of basketball performance was demonstrated by correlating Profile scores with subjective ratings using the Spearman Rho coefficient (12). The average correlation between Profile scores and the average rankings of a board of experts as determined from Fisher's z transformation (13) for twenty observations of inter-collegiate varsity basketball was .785 (significant at the .01 level of confidence). The average validity coefficient for the range of ability found in the classes of grade ten high school boys studied was .884 using the teacher's rating as the criterion and .870 using the Modified Johnson Test as the criterion. Both coefficients were significant ($p = .01$).

These correlations compared favorably with validity coefficients reported for standarized basketball skill tests. The Edgren test (11) yielded a validity coefficient of .77. The Lehsten test (28) demonstrated a validity coefficient of approximately .80. Clarke (6) reported that correlations between the Knox test and subjective ratings of basketball ability were low.

When interpreting the validity coefficients for the Profile

technique five considerations should be made. First, the interjudge reliability as determined by the coefficient of concordance W was barely significant at the .05 level (see table IV). Garrett (16) emphasizes that the correlation between a test and its criterion will be reduced if the criterion scores are unreliable. It would seem probable that the true validity of the Profile Test would be somewhat higher than the achieved values because of the relatively low interjudge reliability.

Second, in the class study the teacher's rating was used as a criterion and it could be expected that subjective ratings of one individual would be less reliable than the average ratings of a board of judges.

Third, there was some evidence to support the hypothesis that the board of experts tended to be somewhat biased by player's past performances and reputations. A t-test for determining the difference between two means for a correlated sample (12) indicated a significant difference ($p = .05$) between the ranges of ranks as determined by the Profile scores and the board of experts. This result was interpreted as an indication that the Profile scores were more sensitive to performance fluctuations than the board of experts. It appeared that the board of experts were reluctant to change pre-conceived impressions.

The hypothesis, that the board of judges were somewhat biased by players' past performances and reputations, could possibly explain why the coefficient of concordance W indicated that the interjudge reliability for the home team was considereably higher than were the coefficients for opponents (see table IV). All of the

home town judges would have known the home team players and could have subconsciously formed preconceived ideas of the relative abilities of the players.

The fourth consideration to be made when interpreting the validity coefficients is that none of the basketball tests mentioned in the review of the literature were validated with samples as highly restricted in range as the present study. Johnson (6) used biserial correlation to demonstrate the validity of his test in dividing subjects into two groups, a "good group" and a "poor group". Knox (6) validated his test by its ability to select successful and unsuccessful candidates for a high school varsity team. In comparison, the Profile technique was used to discriminate among the seven to eleven bets players on intercollegiate and high school varsity teams.

The fifth consideration to be made is the degree of association between the criteria used. Although the correlation between the Profile scores and the teacher's rankings (Rho = .884) and the correlation between the Profile scores and the scores for the Modified Johnson Test (Rho = .870) were very close, analysis of the coefficients of concordance W (see table V p.49) revealed that the association among the measures of basketball performance was low. Five of eight W's were not significant while only three were significant at the .05 level of confidence. It seemed that neither criterion (i.e. the teacher's ranking nor the Modified Johnson Test) contained all the factors implicit in the Profile technique but each one explained a portion of the variance. Therefore it would appear that if a single criterion were available as

a standard for the Profile test, the validity coefficient would probably be higher.

Item Analysis. Item analysis for the intercollegiate varsity scores with time partialled out indicated that each of the three point ratios applied (regular, new, and 422 point systems) yielded discriminating scores on six items (see table VI). The regular point system was the best discriminator on five items while the 422 and new point systems discriminated best on four and two items respectively.

The regular point system was consistently the best discriminator on positive items and the 422 system was the best on negative items. Dean (10) recommends that coaches and teachers emphasize the positive and constructive features of the game rather than develop negative attitudes. It would appear that the regular point system was somewhat better than the other two point ratios applied to the intercollegiate varsity scores.

Field goal shooting discriminated better than any other item yielding a Flanagan index of .82 which supported McCracken's contention (31) that the ultimate question regarding a basketball player is whether he is a scoring threat or not. The offensive rebound item was the next most discriminating with a Flanagan index of .33. Free throws, defensive rebounds and held balls all discriminated well with Flanagan indices of: .27, .27, and .28 respectively.

Fouls discriminated well with a Flanagan index of -.23 using the regular point system and even higher indices with the new and 422 point systems (-.34 and -.38 respectively). The data

indicated that when the amount of time played was partialled out, better players committed fewer fouls than did poorer players at the intercollegiate varsity level of basketball.

Recoveries, yielding a Flanagan index of only .11, did not discriminate well for the intercollegiate players sampled. However, for the high school classes sampled, good discrimination was demonstrated with a Flanagan index of .27 using the regular point system and indices of .32 and .36 using the new and 422 point systems respectively.

Because the difference between recoveries and rebounds was largely a matter of definition, it was decided to determine whether offensive rebounds, an defensive rebounds, recoveries and positive held balls lumped together as a cluster variable designated as take-overs discriminated adequately. Take-overs yielded a Flanagan index of .30 which was higher than would be expected from averaging the Davis indices for each item covered by take-overs and converting to a Flanagan index. The relatively low index of discrimination for recoveries might have indicated that the players sampled from the Western Canada Intercollegiate Athletic Association were, as a group, not very proficient at recovering free balls.

Assists yielded a low index of discrimination (.11) with the regular point system but discriminated well (.38) with the 422 point system. There could be a number of explanations for inadequate discrimination with this item. Perhaps the point system could be improved by other methods than those applied in this study, in such a way that the assist item would be more discriminating but without loss of discrimination on other items. It would be hypothesized that the sample used was generally poorly skilled at

assisting teammates to take unhindered shots. The assist item was found to be the most difficult to objectify and score. Possibly improvement is needed in the item design. Coaches may have employed player specialization with regards to assisting. They might hav instructed players who were not proficient in other skills to assist teammates.

Bad passes and violations did not yield adequate discrimination with the regular point system but did discriminate with the 422 point system. This might suggest that a point ratio could yet be devised which would yield discriminating values on these items without losing discrimination on other items.

Fumbles did not discriminate adequately with any of the three point systems applied and yielded positive discrimination on this negative item with the regular point system. This result is difficult to interpret as it would seem from reviewing the literature and logical reasoning that better players would fumble less than poorer players.

It should be remembered that this study utilized repeated measures on a relatively small, selected sample. It is quite possible that item analysis is simply revealing general lack of proficiency in certain skills within the sampled studied.

Item analysis of the Profile scores derived from study of grade ten high school boys indicated that the new point system discriminated on seven items while the 422 and regular point systems discriminated on six items each. The regular point system was the best discriminator on five items while the 422 and new point systems were best on three and two items respectively. One item was equally discriminating with either the 422 or new point

system.

Although the new point system discriminated on one more item than the regular point system, only one of the items for which it was the best selector discriminated at the criterion level of .20 on the Flanagan scale, while four of the items for which the regular system was the best selectors discriminated at the criterion level of .20 or better. The new point system did not satisfy some of the arguments presented in the development of curricular validity. Therefore, it would be irrational to accept the new point system over the 422 or regular point systems unless there was a marked superiority in item validity and reliability. As indicated in table XIII p. 59, the new point system was not more reliable than the regular or 422 point system.

Almost equal discriminating power was shown by the 422 and regular point ratios. Each one discriminated on six items. The regular system was the better selector on six items while the 422 point system was the better selector on five items. Both the 422 and the regular point systems discriminated well on all the positive items except assists. Assists yielded negative discrimination with both the 422 and the regular point ratios.

Apparently, the less skilled players tended to pass the ball to the more skilled players when the ball was in scoring range. It must be remembered that these results were taken from a class situation in which groups had no previous opportunity to practice or play together as a team. Therefore it would be reasonable to assume that the players had not developed the teamwork expected for groups which usually play together as a team.

None of the negative items discriminated well. Bad passes, using the regular point system, approached the criterion level with a Flanagan index of .18 but violations yielded a positive index of .25 for this negative item. With all three point ratios, fouls produced large positive indices of discrimination. This would appear to indicate that the better players foul more than the poorer players. This is in direct contrast to the results from the intercollegiate varsity study which indicated that the better players foul less than less skillful players.

At the grade ten class level of basketball, it was observed that the better players were more aggressive. They got more rebounds, recoveries and held balls than the less agressive players but made more mistakes in doing so. They were probably in better defensive position to stop their opponents from advancing the ball and thus were more vulnerable to fouls than players who were less aggressive and who probably did not defend as closely. A reasonable explanation of the difference between the indices of discrimination for the fouls item for intercollegiate varsity players and grade ten boys would be that it takes years of practise to develop players so that they can defend closely without fouling. Intercollegiate varsity players will have had years of practice and could be expected to defend well with fewer errors than the high school boys.

Th inability of the negative items to demonstrate adequate discrimination could be explained largely by the number of times players had the ball. From the number of shots taken and rebounds, recoveries and held balls secured, it was reasonable to infer that

the better players had the ball more often than the poorer players. Therefore, to get an accurate analysis of the discrimination of these negative items, one would have to partial out the number of times an individual had the ball.

Reliability. The reliability of the Profile technique can be considered in relation to three uses to which the methods could be applied. First, as a measure of basketball performance on single occassions, the average reliability coefficients as determined through the use of Fisher's z transformation (13) were .596 and .590 for intercollegiate and high school varsity basketball respectively. Neither reliability coefficient was significant at the critical value ($p = .01$).

Second, as a measure of basketball ability over a period of four or more games, and reliability coefficients was significant ($p = .01$). The average coefficients found by using the Fisher z transformation (13) were: .750, .795 and .838 over a period of four, five and six games respectively.

The third use for which the Profile technique could be applied is as a test. Average reliability coefficients were determined through Fisher's z transformation (13) and corrected for restriction of range by the procedures outlined by Guilford (19). Reliability coefficients for the range of ability found in the classes of grade ten boys sampled were: .854, .842 and .869 employing the regular, new and 422 point system respectively.

All three coefficients were significant at the .01 level of confidence and coefficients computed by applying the regular and new point systems were above the .85 value recommended by Scott

(37) for test reliability.

It was perceived that the reliability of the Profile technique as a measure of basketball ability over a period of four or more games for intercollegiate varsity basketball and as a test for grade ten high school boys was significant at the critical value (p = .01). However, as a measure of basketball performance in a game situation for both intercollegiate and high school varsity basketball, the reliability coefficients were considerably lower than the level of confidence (p = .01) set as the criterion value.

In search for a reasonable explanation of the low single game reliability, three variables were considered. First, the variable amount of playing time may have caused the low reliability. It was hypothesized that small chance errors in the scores of substitutes who played for relatively short periods of time were being magnified when the results were equated for time. The sum of squares of differences between paired ranks for the upper half (regulars) and the lower half (substitutes) of the subjects according to time played did not indicate any consistent difference between groups (see table XI p. 57). Therefore it would appear that the principle variable which depressed the reliability was not the time factor.

Second, it was hypothesized that there were considerable intraindividual performance fluctuations caused by differential fatigue, warm-up and speed factors. To test this hypothesis, the ranges of ranks for each player on the University of Alberta (Edmonton) team were determined for both the experts' ratings and Profile scores. The mean range of ranks as determined by the

Profile scores was 7.56 while the mean range of ranks as determined by the experts' ratings was 5.28. These means were significantly different ($p = .05$). This result was interpreted as an indication that there were large intraindividual performance fluctuations and that the Profile technique was more sensitive to these fluctuations than was the board of judges.

If performance fluctuations are as great and as frequent as were indicated by the range of player Profile ranks and the low single game reliabilities found and if the Profile technique is as sensitive to these changes as it appears, then Profile technique could be a valuable aid to the coach in assessing player performance from quarter to quarter or half to half. One of the characteristics a coach might wish to consider when selecting his team could be reliability of performance. The Profile technique could give the coach an indication of the player's intragame and intergame reliability.

The third possible explanation for the low single game reliability is that the variations in teammates and opponents caused by substitutions could have lowered the reliability. Two time studies were undertaken to determine the amount of time a player must play to obtain reliable results. However, the results of these studies shed more light on the effect of substitutions than on the time variable.

In time study I the reliability improved with increased playing time but in time study II the reliability decreaded as playing time was increaded. Ferguson (12:283) states, "From a theoretical point of view a test may be made as reliable as we like by increasing its length." If the assumptions underlying the

Spearman-Brown formula were met, one could have expected reliability to improve with increased time. As Guilford (19) has stated, "An estimate by the Spearman-Brown formula is probably conservative, because it tends to be an underestimate."

It would appear that the assumptions underlying the valid application of the Spearman-Brown formula must not have been met in the time studies observed. The fundamental assumption for the valid use of the Spearman-Brown formula is comparability of halves (19). Comparability of halves implies that content must be similar, boredom and fatigue must be minimal, and there must be an equal range of difficulty (16).

Content in the controlled scrimmages was identical and there was little evidence of boredom or differential fatigue within the duration of the studies. The most probable explanation for the low reliability was that equal difficulty must not have been involved from one five minute period to another. In time study I the experimental team was kept constant but one or two substitutes were made for their opponents in each five minute period. In time study II one substitute was made on both the experimental team and that of the opponents.

Indications were that even one substitute upset the reliability considerably. In the controlled scrimmages all players on the roster were used while in game situations only the best eight or nine played. Therefore, the range of ability sampled in the time studied was greater than the range sampled in the intercollegiate games. In the time studies, a first string player would often be replaced by a third string substitute. This would change the ranking of abilities on the team for which the substitute

was made and could even reverse the ranking for the opponent of the substitute. Instead of playing opposite one of the most difficult players to guard defensively and to outmaneuver offensively, he would now be set against one of the least skilled opponents.

It was evident that one of the most important variables to control when employing the Profile technique as a test was the selection of teammates and opponents. The variation in teammates and opponents would be the most probable reason for the relatively low reliability for single games since a player frequently played opposite different opponents in the second half than in the first. The substantially higher reliability in the high school class study tended to support this hypothesis. In the class study: time, teammates and opponents were controlled.

Because of the uncontrolled variable of teammates and opponent placement little can be concluded regarding the amount of time a player must play to obtain reliable results.

It is expected that reliability could be improved by two methods. The first method would be to control the selection of teammates and opponents. If only one testing opportunity was available, perhaps the best way to assign players to positions would be to match them according to some form of standardized pre-testing as they were in the present class study. On the other hand, if a reasonably long testing period or a number of occassions were available, the results would indicate that it would be advantageous to rotate players at standard intervals so that they would play against a variety of opponents. This procedure could be

expected to produce a regression effect towards true scores.

Of course these procedures are not practical for the evaluation of performance in actual game situations. The results indicated that over a period of four or more games that there was a sufficient regression towards true scores as the result of repeated measures and a sufficient variety of opponents and teammates. However, it seems unlikely that single game reliability could be improved significantly.

Ferguson (12:288) has implied that low reliability does not negate the legitimate use of an instrument for the evaluation of performance:

> Low reliability does not necessarily invalidate a a technique as a device for drawing valid inferences. Low reliability may be compensated for by increase in sample size. An unreliable technique used with a small sample is, however, capable of detecting gross differences only.

According to Ferguson's discussion, it could be concluded that the Profile technique used without careful selection of teammates and opponents on a single occassion or on less than four occassions would be valid for detecting gross differences only. On the other hand, if selection and placement of participants is cautiously regulated or if repeated measures are provided to minimize the selective variable through the regression effect, more delicate inferences may be drawn.

<u>Objectivity</u>. The Profile technique exhibited a high degree of objectivity for the managers' scores when correlated with the author's scores. Average objectivity coefficients as computed through Fisher's z transformation (13) were .910 for both the high school and the university managers' scores (significant at the

.001 level of confidence).

Three variables were perceived to affect the objectivity coefficients. The first variable was player familiarization. Both the university manager and the high school manager obtained scores that were more objective for their own teams than for visiting teams. With the home team, and observer usually knew the players' numerals and could accurately record scores even if the numerals were not in view. With unfamiliar players, obscured numerals sometimes led to incomplete player Profile point accumulation (see table XV, p. 61).

The second variable that affected objectivity coefficients was the type of numerals on the players' uniforms. Large block numerals in a contrasting colour to the body of the uniform facilitated the taking of objective scores. It should be noted that the lowest objectivity coefficient obtained was for the University of Manitoba team (see table XV, p. 61). This team wore uniforms with yellow numerals on a white background.

The third variable which influenced objectivity coefficients was the amount of skill an observer had with the Profile technique. Facility with the Profile technique tended to improve over the first three exposures but remained fairly constant thereafter.

As pointed out earlier (see tables III, p. 40 and X, p. 56) both the reliability and validity coefficients for the University of Alberta (Edmonton) team were larger than those computed for their opponents. A reasonable explanation was that the higher objectivity coefficients for familiar teams indicated a greater degree of accuracy in scoring these teams in comparison with unfamiliar teams. This would in turn affect both validity and reliability.

Compared to the degree of agreement among the board of experts the Profile technique manifested superior objectivity (see tables IV, p. 42 and XV, p. 61). The board of experts, with their considerable experience and background and experience demonstrated agreement that was barely acceptable at the .05 level of significance after observing complete games. It would appear that subjective grading in physical education classes, by a single person, generally less experienced than the members of the board of experts and usually with less time for observation, is a dubious practice.

The high objectivity coefficients objectivity coefficients obtained (see table XV, p. 61) indicated that university and high school students were able to take accurage Profile data without extensive training or experience in basketball. Normal intelligence, an interest in the game, and a desire to do a good job appeared to be the prime requisites for the Profile observer. However, it was evident that familiarity with the players faciliteted the establishment of accurate data. Objectivity was observed to improve with practice over the first three applications, after which scores tended to remain at a fairly constant level as evidenced by the coefficients which remained in the range significant at the .001 level for familiar teams and at the .01 level for unfamiliar teams.

CHAPTER V

SUMMARY AND CONCLUSIONS

Summary. The primary purpose of this study was to develop an objective instrument, which was valid, reliable and economical in terms of cost of the instrument and time, for the evaluation of competitive basketball performance. Secondary purposes of the study were (1) to determine the length of performance necessary to obtain valid and reliable evaluation; (2) to determine if the instrument was suitable for evaluation of: intercollegiate varsity basketball, and second year high school class basketball; (3) to determine if the instrument could be used as a grading device; (4) to determine the amount of training necessary for an observer to obtain accurate results; and (5) to determine whether reliability and item validity could be improved by changing the point ratio of the Profile technique.

Three sets of observations were made. The first set included twenty observations of intercollegiate varsity basketball competition. The second set of data was taken from senior high school varsity basketball and the third set of observations were taken from two classes of grade ten boys. The principal data gathering device used was the "Basketball Profile Statistics Chart".

Three dimensions of curricular validity were shown. First, it was demonstrated that the material covered by the Profile technique was important to successful performance in basketball competition. Second, it was shown that the various test items were scored with a logical ratio of emphasis. Third, it was shown that the Profile Test closely simulates a real situation.

Curricular validity was supported by correlating Profile scores for the intercollegiate games with the average rankings of a board of judges using Spearman's coefficient of rank correlation Rho. Forty-four grade ten boys were matched according to the results of the Modified Johnson Basketball Test for Profile testing. Profile scores for the grade ten classes were correlated with the teacher's rankings and with the results of the Modified Johnson Test by means of the Spearman Rho technique.

The discriminating power of each Profile item was determined by both the Davis and Flanagan techniques.

Reliability for all Profile scores was determined by correlating first half scores with second half scores using the Spearman Rho coefficient and estimating the reliability of the whole test by the Spearman-Brown formula.

Objectivity of the Profile technique was determined by correlating the author's scores with the results obtained independently by the managers of the University of Alberta (Edmonton) team and with the scores obtained independently by the manager of the Strathcona senior high school varsity team. Because so many coefficients were calculated it was necessary to group data by means of the Fisher z transformation.

The validity coefficient for intercollegiate varsity basketball using the average rankings of a board of experts was .785 which was significant ($p = .01$). For the range of ability found in the classes of grade ten boys studied, the validity coefficients using the teachers rating and results of the Modified Johnson Test as criteria were .884 and .870 respectively. Both coefficients

were significant at the .01 level of confidence.

Item analysis for the intercollegiate varsity scores with time partialled out, indicated that each of the three point ratios applied yielded discriminating scores on six items. The regular point system was the best discriminator on five items, while the 422 and new point systems discriminated best on four and two items respectively. The regular point system was consistently the best discriminator on positive items and the 422 point system was the best on negative items.

Item analysis of the Profile scores derived from study of grade ten high school boys indicated that the new point system discriminated on seven items while the 422 and regular point system discriminated on six items each. The regular point ratio was the best on five items while the 422 and new point ratios were best on three and two items respectively. One item was equally discriminating with either the 422 or new point system.

Field shots, free throws, offensive rebounds, defensive rebounds, and held balls discriminated well for both the intercollegiate varsity players and grade ten class players sampled. Recoveries discriminated well for the intercollegiate varsity sample. Assists discriminated well using the 422 point system for the intercollegiate sample but did not with the regular point system. Negative discrimination resulted for the assist item using the regular point ratio with the grade ten class scores.

Fouls discriminated well for the intercollegiate varsity sample but produced positive discrimination for this negative item with the high school class. Other than fouls for the intercollegiate

varsity sample, none of the negative items discriminated adequately with the regular point system.

The reliability of the Profile technique can be considered in relation to three uses to which the method could be applied. First, as a measure of basketball performance on single occassions, the reliability coefficients were .596 and .590 for intercollegiate and high school varsity basketball respectively. Neither reliability coefficient was significant at the critical value (p = .01). Second, as a measure of basketball ability over a period of four or more games, the reliability coefficient was significant (p = .01). Coefficients for four, five and six games were: .750, .795 and .838 respectively. The third use for which the Profile technique could be applied is as a test. Reliability coefficients corrected for restriction of range for the Profile Test as administered to the grade ten boys sampled were: .845, .842 and .869 employing the regular, new and 422 point systems respectively. All three coefficients were significant at the .01 level of confidence.

Objectivity coefficients between the author's Profile scores and those taken independently by both high school and university managers were .910 (significant at the .001 level of confidence). Both the university and the high school managers were able to obtain scores that were more objective for their own teams than for opponents. Both managers were able to obtain scores that were objective at the .01 level of confidence for both home and visiting teams on the third application of the Profile technique.

<u>Conclusions</u>. Within the limitations of the statistical procedures employed and the populations investigated, the following

conclusions appear to be justified.

The Profile tecnique was demonstrated to be a highly objective instrument for the evaluation of basketball performance. Objectivity coefficients significant at the .001 level of confidence indicated that university and high school students were able to take accurate Profile data without extensive training or experience in basketball. The results illustrated that the high school manager participating in the study learned to use the Profile technique accurately for familiar players after one practice application and for unfamiliar players, after two practice trials.

The university manager participating in the study was able, to obtain accurate scores for the home team on the first application but required two practice occassions before accurate scores could be obtained for visting teams. It was concluded that two practice trials provided sufficient training for the Profile observer.

Profile evaluation was found to be economical in cost and required two referees, one Profile observer and one spotter to administer.

When subjects were carefully matched, valid evaluation of grade ten boys was made in twenty minutes. When players were not carefully matched only gross inferences could legitimately be drawn from performance in less than four games.

The validity of the Profile technique was established by the curricular method. As a measure of intercollegiate basketball in game situations, the validity of the Profile technique was given further support by significant correlations (p = .01) with the average rankings of a board of experts. As a test for grade ten

boys, the validity of the Profile method, with the teacher's ratings and the results of the Modified Johnson Test as criteria, was upheld by correlations significant at the .01 level of confidence.

As a test for grade ten high school boys, the Profile technique was found to be reliable at the .01 level of confidence. Profile evaluation of basketball ability for intercollegiate varsity players was found to be reliable ($p = .01$) over a period of four or more games. For evaluation of high school varsity and intercollegiate varsity basketball performances in single game situations, Profile scores were not reliable at the .01 level of confidence set as the criterion.

For intercollegiate varsity basketball and for second year high school class basketball, the Profile method was shown to be an appropriate measuring device. Insufficient data was presented to make any firm conclusions regarding the suitability of the Profile technique as a measure of high school varsity basketball.

The results indicated that the Profile technique met the requirements advanced by Clarke (6) of a "good test". However, it should be cautioned that scores are relative to the level of performance within the group being tested. Therefore, it can be concluded that the Profile test is a suitable instrument for grading students by instructors who believe in relative grading.

Reliability was not improved significantly by the two changes in the point ratio applied in the present study. There was some evidence that changes could be made in the point ratio to improve item validity but neither of the alternate point systems applied in the present study demonstrated marked improvement over the regular

point system. However, there could be some justification in using the 422 point system for high school class basketball.

Field shots, free throws, offensive rebounds, defensive rebounds and held balls were discriminating items. The inability of the negative items to show consistent discrimination was probably caused by the unequal number of times that each player had the ball.

The Profile technique can be legitimately used for three purposes: (1) as a test, (2) as a measure of basketball ability over a period of four or more games and (3) for drawing gross inferences regarding basketball performance in single game situations.

The Profile Test has three fundamental advantages over conventional methods of basketball evaluation. First, it measures a realistic competitive situation, Second, it is diagnostic in that the source of errors is clearly demonstrated. Third, because it is realistic, diagnostic, and fun for most subjects, students and players are strongly motivated by it's use.

Recommendations. Much work is yet to be done to collaborate the present results, to improve the Profile technique, and to investigate inconclusive areas of the present study. Listed below are a number of problems which require special attention:

1. Because the present time studies were confounded by uncontrolled assignment of subjects and substitutions, further study is necessary to determine the minumum amount of time required for valid and reliable evaluation under controlled conditions.

2. The validity of the Profile technique as a measure of high school varsity basketball has yet to be shown.

3. Further study is necessary to determine why the recovery and assist items did not discriminate adequately for intercollegiate

varsity basketball.

4. The hypothesis that the inability of the negatively scored Profile items to demonstrate adequate discrimination could be explained by the number of times players had the ball should be investigated. To get an accurate analysis of the discrimination of negatively scored items, one would have to partial out the number of times an individual had the ball.

5. The hypothesis that variation in teammates and opponents caused by substitutions was the primary variable which lowered reliability, needs to be collaborated.

6. There is a need for study of various methods of assigning subjects to positions.

7. Study of various populations is necessary to distinguish between results caused by factors inherent in Profile testing and population variance.

8. Films could be used to advantage in study of the Profile technique. First, films could be used to determine the source of errors. Second, films could be used to provide the board of experts with a careful, detailed observation of the test situation which could be expected to result in a more reliable criterion. Third, various techniques of basketball evaluation could be observed by the use of films to determine which method best reduces errors in scoring.

BIBLOGRAPHY

1. Anderson, Forrest, Basketball Techniques Illustrated. New York: A. S. Barnes Compnay, 1952 95 pp.

2. Basketball Instructor's Guide, Chicago: The Athletic Institute, 1952 72 pp.

3. Bovard, J. F., Cozens, F. W., and Hagman, E. P., Tests and Measurements in Physical Education. Philadelpia: W. B. Saunders Company, 1949.

4. Brace, David K., "Testing Basketball Technique", American Physical Education Review, XXIX (March, 1924), 159-165.

5. Case, Wayne F., "Statistical Chart Data-Advantages and Disadvantages", Athletic Journal, XLV (October, 1964), 12.

6. Clarke, H. Harrison, Application of Measurement to Health and Physical Education. Englewood Cliffs, N. J.: Prentice Hall, Inc., 1959. 528 pp.

7. Colver, Robert M., "Estimating Item Indices By Nomographs", Psychometrika, 24: 179-185, June 1959.

8. Cratty, Bryant J., Movement Behavior and Motor Learning. Philadelphia: Lea and Febiger, 1964. 332 pp.

9. Davis, Frederick B., Item-Analysis Data: Their Computation, Interpretation, and Use in Test Construction. Harvard Education Papers No. 2, Graduate School of Education, Harvard University 1946.

10. Dean, Everett S., Progressive Basketball Methods and Philosophy. Englewood Cliffs, N. J.: Prentice-hall, Inc., 1946. 271 pp.

11. Edgren, H. D., "An experiment in Testing of Ability and Progress in Basketball", Research Quarterly, 3:159-171, March, 1932.

12. Ferguson, George A., Statistical Analysis in Psychology and Education. Toronto: McGraw-Hill Company, Inc., 1959, 347 pp.

13. Fisher, R. A., Statistical Methods for Research Workers. London: Oliver and Boyd Company, 1938.

14. Flanagan, J. C., "General Considerations in the Selection of Test Items and a Short Method of Estimating the Product-Moment Coefficient from Data at the Tails of the Distribution", Journal of Educational Psychology, 30:674, 1939.

15. Fuoss, Don, "Grading Your Football Films", Scholastic Coach, 34:62, January, 1965.

16. Garrett, Henry E., Statistics in Psychology and Education. Toronto: Longmans, Green Company, 1953. 460 pp.

17. Gates, Paul L., "Scouting with Radio and Tape Recorder", Scholastic Coach, 34:34, January, 1965,

18. Glassow, Ruth B., Colvin, with Valerie, and Schwartz, Marguerite M., "Studies in Measuring Basketball Playing Ability of College Women", Research Quarterly, 9:60-68, December, 1938.

19. Guilford, J. P., Fundamental Statistics in Psychology and Education. Toronto: McGraw-Hill Book Company, 1956.

20. Handler, Fred D., "Basketball Scouting", Athletic Journal, XLIV (October, 1963), 32.

21. Henry, Franklin, and Smith Leon E., "Simultaneous vs. Separate Bilateral Muscular Contrations in Relation to Neural Overflow Theory and Neuromotor Specificity", Research Quarterly, 32: 42-46, March, 1961.

22. Holman, Nat, Winning Basketball. New York: Charles Scribner's Sons, 1932.

23. Hunter, Kenneth A., "Basketball Scouting", Athletic Journal, XLIV (November, 1965), 16.

24. Kelley, T. L., "The Selection of Upper and Lower Groups for the Validation of Test Items", Journal of Educational Psychology, 30: 17-24, 1939.

25. Kenyon, Gerald, "Why Measurement in Physical Education", Journal of the Canadian Association of Physical and Health Education and Recreation, XXIII (March 1957), 17.

26. Kowalk, Clayton J., "The Fundamental Outloook", Athletic Journal, XLVI (September, 1965), 18.

27. Larson, L. A., and Yokum, R. D., Measurement and Evaluation in Physical, Health and Recreation Education. St. Louis: The C. V. Mosby Company. 1951.

28. Lehsten, N., "A Measurement of Basketball Skills in High School Boys", The Physical Educator, 5:103-109, December, 1948.

29. Longfellow, Floyd Jr., "Fast Break Analysis Chart", Athletic Journal, XVI (November, 1965), 16.

30. Masin, Herman L., editor, "Coaching With Videotape", Scholastic Coach, 35: 36, January, 1966.

31. McCracken, Branch, Indiana Basketball. New York: Prentice Hall Inc., 1955, 207 pp.

32. McVay, John E., "Utilization of Game Film", Athletic Journal, XLIV (February, 1964), 14.

33. Mortimer, Elizabeth M., "Basketball Shooting", Research Quarterly, 22:234-243, May, 1951.

34. Noch, George D., "Drills for Successful Rebounding", Athletic Journal, XLIV (November, 1963), 42.

35. Olds E. G., "Distribution of Sums of Squares of Rank Differences for Small N's", Annals of Mathematical Statistics, 9:145, 1938.

36. Oxford, Harry J.,"All Fundamentals the Thirty-Minute Way", Athletic Journal, XLV (December, 1964), 40.

37. Scott, M. Gladys, editor, Research Methods in Health Physical Education Recreation. Washington, D.C.: American Association for Health, Physical Education and Recreation, 1959. 536 pp.

38. Scott, M. Gladys, and French, E., Measurement and Evaluation in Physical Education. Dubuque: William C. Brown Company, 1959.

39. Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences. Toronto: McGraw-Hill Book Company Inc., 1956. 312 pp.

40. Summers, Art, "Correct Basketball Statistics Using a Tape Recorder", Athletic Journal, XLV (September, 1964), 31.

41. Trump, Cliff, "Meaningful Grading", Scholastic Coach, 35:44, January, 1966.

42. Wooden, John R., Practical Modern Basketball. New York: The Ronald Press Company, 1966. 418 pp.

43. Young, Genevieve, and Moser, Helen, "A Short Battery of Tests to Measure Playing Ability in Women's Basketball", Research Quarterly, 5:3-23, May 1934.

APPENDIX A

"BASKETBALL PROFILE" STATISTICS CHART

"BASKETBALL PROFILE" STATISTICS CHART

Date__________ School____________ VS ______________ at_________ Score_______ Half

Negative Points			Positive Points		
BAD PASSES	VIOLATION	FUMBLE	± HELD BALL	RECOVERY	ASSISTS

SCHOOL
SHOTS Att.______ Made______ Pct.______

REBOUNDS

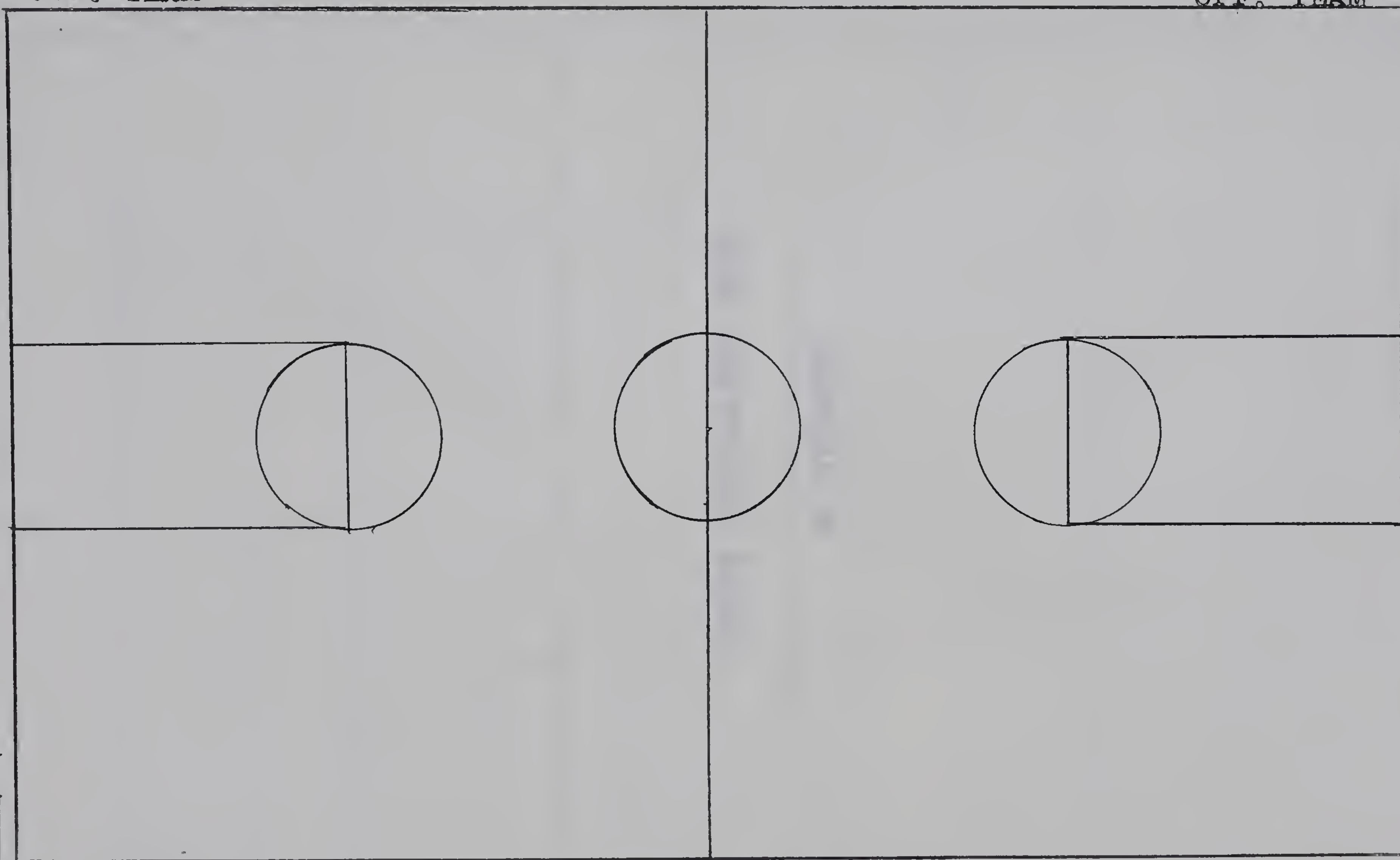
OFFENSIVE

DEFENSIVE

FOULS

OFF. TEAM

OFF. TEAM



FREE THROWS

SCHOOLS
SHOTS Att.______ Made______ Pct.______

REBOUNDS

OFFENSIVE

DEFENSIVE

FOULS

APPENDIX B

DATA COMPUTATION SHEET

vs AT DATE OBSERVER

TEAM		FIELD SHOTS			FREE THROW			REBOUNDS			POSITIVE PTS				NEGATIVE PTS					FOULS		TOTAL
No.	NAME	SA	SM	PTS	SA	SM	PTS	OFF	DEF	PTS	HELD BALL	REC	AST	PTS	HELD BALL	BAD PASS	VIO	FUM	PTS	PF	PTS	PTS
	TOTALS																					
TEAM	NAME																					
	TOTALS																					

APPENDIX C

MODIFIED JOHNSTON BASKETBALL TEST

I. Field Goal Speed Test.

<u>Equipment Neccessary</u>- stop watch.

1. One basketball.
2. One basketball backboard and goal.

<u>Directions</u>-

1. Shoot as many baskets as possible in 30 seconds.
2. You may shoot from any point on the floor.
3. Signal- "Ready, Go!".

<u>Scoring</u>- one point for each basket made in 30 seconds.

II. Agility Dribble.

<u>Equipment Necessary</u>- stop watch.

1. One basketball.
2. Starting line 3 feet long.
3. Four chairs, the first one 12 feet from the starting line and each additional chair 6 feet apart in a straight line.

<u>Directions</u>-

1. From the starting line on the signal, "Ready, Go!" dribble to either side of the first chair and weave in and out of each chair dribbling with the outside hand.
2. When you get to the end chair go all the way around it and weave through the chairs on the way back.
3. When you return to the first chair, go all the way around it and continue weaving until the whistle signal to stop

NB- After the first round do not go back to the starting line.

<u>Scoring</u>-

1. Number of chairs passed in 30 seconds minus number of error zones.
2. Error zones are those in which the subject makes a fumble, violation or dribbles with the inside hand.

NB- count end chair as 1 only.

III. Wall Pass.

<u>Equipment Necessary</u>- stop watch.

1. Line 5 feet from an unobstructed wall space at least 6 feet wide and 6 feet high.
2. One basketball.

Directions-

1. From the starting line pass the ball with a two hand chest pass against the wall so that it's rebound can be caught on the fly.
2. Pass and catch the ball as rapidly as possible in the 30 seconds.
3. Start on the signal, "Ready, Go!" and continue to the whistle signal to stop.
4. If the ball goes astray, retrieve it and start again from behind the starting line.

Scoring-

1. One point is awarded for each catch made behind the restraining line.
2. The ball must be caught cleanly and not trapped against the body.
3. Do not count catches in front of the restraining line or catches of the rebound of throws made from infront of the line.

APPENDIX D

TABLE XIX

DETAILED LIST OF
PROFILE VALIDITY COEFFICIENTS FOR
INTERCOLLEGIATE VARSITY BASKETBALL

Team	N	Rho	Significance
U. of Alberta (Edmonton)	11	.43	not significant
	8	.95	.01
	9	.89	.01
	9	.16	not significant
	9	.97	.01
	10	.64	.05
U. of Saskatchewan	7	.44	not significant
	7	.86	.05
U. of Manitoba	10	.67	.05
	9	.84	.01
U. of Alberta (Calgary)	8	.79	.05
	7	.73	.05
	9	.74	.05
	10	.89	.01
U. of Windsor	10	.94	.01
	10	.79	.01
Acadia University	7	.89	.01
	9	.79	.01
Carleton University	8	.76	.05
	6	-.09	not significant

APPENDIX E

TABLE XX

DETAILED LIST OF VALIDITY COEFFICIENTS
OF A TWENTY MINUTE PROFILE BASKETBALL TEST
FOR GRADE TEN BOYS
USING THE TEACHER'S RATING AS CRITERION

Class	Group	Rho (small range)	Rho (large range)
3 and 4	I. green	-.625	.400
	I. gold	.700	.926
	II. green	.500	.876
	II. gold	.800	.951
5 and 6	I. green	.900	.975
	I. gold	.600	.901
	II. green	.600	.901
	II. gold	.075	.772

APPENDIX E

TABLE XXI

DETAILED LIST OF VALIDITY COEFFICIENTS
OF A TWENTY MINUTE PROFILE BASKETBALL TEST
FOR GRADE TEN BOYS
USING THE RESULTS
OF THE MODIFIED JOHNSON TEST AS CRITERION

Class	Group	Rho (small range)	Rho (large range)
3 and 4	I. green	-.325	.673
	I. gold	.500	.876
	II. green	.700	.926
	II. gold	.800	.951
5 and 6	I. green	.900	.975
	I. gold	.700	.926
	II. green	.200	.802
	II. gold	-.525	.623

APPENDIX F

TABLE XXII

DETAILED LIST OF RELIABILITY COEFFICIENTS FOR PROFILE EVALUATION OF SINGLE INTERCOLLEGIATE VARSITY GAMES

Team	N	Rho	Significance
U. of Alberta (Edmonton)	9	-.02	not significant
	8	.73	.05
	8	.61	not significant
	10	.52	not significant
	9	.73	.05
	8	.82	.05
U. of Saskatchewan	7	.53	not significant
	7	.31	not significant
U. of Manitoba	7	-.70	not significant
	8	.67	.05
U. of Alberta (Calgary)	8	.49	not significant
	8	.83	.05
	9	.54	not significant
	9	.50	not significant
U. of Windsor	8	.55	not significant
	9	.78	.05
Acadia University	6	.85	.05
	6	.87	.05
Carleton University	7	.18	not significant
	6	.34	not significant

APPENDIX G

FORMULAE USED

1. Spearman's Coefficient of Rank Correlation Rho. Rho is a measure of association which requires that both variables be measured in at least the ordinal scale so that the objects or individuals under study may be ranked in two ordered series (39). Ferguson (12) presents the Rho formula as follows:

$$P = 1 - \frac{6 \sum d^2}{N(N^2 - 1)}$$

Where:

$N =$ the number of subjects or objects studied.

$\sum d^2 =$ the sum of the differences between paired ranks.

2. The Kendall Coefficient of Concordance W. When we have k sets of rankings, we may determine the association among them by using the Kendall coefficient of concordance W. Such a measure may be particularly useful in studies of interjudge or intertest reliability (39).

$$W = \frac{12\ S}{m^2\ (N^3 - N)}$$

Where:

$m =$ number of judges

$N =$ number of subjects

$$S = \sum \left(R_j - \frac{\sum R_j}{N}\right)^2$$

Where:

R_j represents the rank sum of the j th individual.

S represents the sum of squares of rank sums for N individuals.

3. Correction for Attenuation. The correlation between a test and its criterion will be reduced if either the test scores or the criterion are unreliable. Correction for attenuation is used to estimate the correlation between true scores in two variables (16).

$$r_{\infty\infty} = \frac{r_{12}}{\sqrt{r_{1I} \times r_{2II}}}$$

Where:

$r_{\infty\infty}$ = correlation between true scores in tests one and two.

r_{12} = correlation between tests one and two.

r_{1I} = reliability coefficient of test one.

r_{2II} = reliability coefficient of test two.

4. The T-Scale. "T" scores normalize the frequency distribution of standard scores (16). The T-Scale has a mean of 50, a standard deviation of 10, and ranges from 0 to 100. Most scores fall between 15 and 85.

Steps:

1. Calculate frequency of each interval.

2. Do a cumulative frequency for the top score down.

   (a) Cumulative frequency for top score = N.
   (b) Cumulative frequency for second top score = N - f of top score - etc.

3. Calculate cumulative frequency to mid point of each frequency.

4. Take cumulative f to mid point and convert to percent by taking:

$$\text{PERCENT} = \frac{\text{CUMULATIVE f TO MID POINT}}{\text{N}} \times 100$$

5. Percents can be turned into "T" scores by table XXII below.

TABLE XXIII

CALCULATION OF "T" SCORES FROM CUMULATIVE FREQUENCIES TO MID POINTS IN TERMS OF PERCENT

Percent	T-score	Percent	T-score
.023	15	69.150	55
.130	20	84.130	60
.620	25	93.320	65
2.280	30	97.720	70
6.680	35	99.380	75
15.870	40	99.865	80
30.850	45	99.977	85
50.000	50	99.997	90

5. Spearman-Brown Formula. With the split half method of reliability, the whole test is more reliable than either half and there is generally an increase in reliability with increased length of the test (19). The Spearman-Brown formula can be used to estimate the effect of lengthening or repeating a test (16).

$$r_{nn} = \frac{n\, r_{1I}}{1 + (N-1)\, r_{1I}}$$

Where:

r_{nn} = the estimate of the correlation between n alternate forms or (the mean of n forms against the mean of n other forms).

r_{1I} = reliability of the present test.

6. Correction For Restriction of Range. The reliability of a test is affected by the variability of the group (16). Often it is desired to know the self-correlation of a test in a larger group since it is probably more stable from population to population (19). The coefficient of validity is almost invariably smaller in a restricted group than it would be in a relatively

unrestricted group. If one wishes to avoid wrong conclusions when a substantial amount of selection has been made, one should apply correction procedures (19).

$$r_{11} = 1 - \frac{S.D._s^{\,2}\ (1 - r_{ss})}{S.D._1^{\,2}}$$

<u>Where</u>:

r_{11} = correlation in larger range.

r_{ss} = correlation in smaller range.

$S.D._s$ = standard deviation in smaller range.

$S.D._1$ = standard deviation in larger range.